bioinformatics: Gordon D Brown
POLYA=Yes.

FEATURES Location/Qualifiers
source 1. .212
/organism="Oncorhynchus mykiss"
/mol_type="mRNA"
/strain="Karl Vierke"
/db_xref="taxon:8022"
/clone_lib="Oncorhynchus mykiss reproductive"
/note="Vector: pcr4topo; Library Creator: Kristian R von Schalburg ; Rainbow trout tissue contributors: Mountain Trout Sales (Sooke, B.C.)"

ORIGIN

Query Match 84.0%; Score 16.8; DB 14; Length 212;
Best Local Similarity 90.0%; Pred. No. 2.2e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CTGGACAGGTTAGGGCTTTG 20
||||||||||| ||| |||||
Db 111 CTGGACAGGTGAGGACTTTG 92

RESULT 15
BB306076
LOCUS BB306076 259 bp mRNA linear EST 10-JUL-2000
DEFINITION BB306076 RIKEN full-length enriched, adult male corpora quadrigemina Mus musculus cDNA clone B230206M01 3', mRNA sequence.
ACCESSION BB306076
VERSION BB306076.1 GI:9006781
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 259)
AUTHORS Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N., Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y., Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tominaga,N., Toya,T., Tsunoda,Y., Watahiki,A., Watanabe,S., Yamamura,T., Yamanaka,I., Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE RIKEN Mouse ESTs (Konno,H., et al.)
JOURNAL Unpublished (2000)
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J., Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)

Please visit our web site (http://genome.rtc.riken.go.jp) for further details.

FEATURES Location/Qualifiers
source 1. .259
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="B230206M01"
/sex="male"
/tissue_type="corpora quadrigemina"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, adult male corpora quadrigemina"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3']. cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 259;
Best Local Similarity 90.0%; Pred. No. 2.2e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CTGGACAGGTTAGGGCTTTG 20
|| ||||||||| |||||||
Db 72 CTTGACAGGTTATGGCTTTG 91

Search completed: February 20, 2004, 01:41:36
Job time : 137.422 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 21:32:00 ; Search time 3.2454 Seconds
(without alignments)
3419.919 Million cell updates/sec

Title: US-09-939-853A-140
Perfect score: 20
Sequence: 1 ctggacaggttagggctttg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		15.8	79.0	99916	4	US-09-816-095-3	Sequence 3, Appli
c 2		15.2	76.0	303	4	US-09-489-039A-4299	Sequence 4299, Ap
3		15.2	76.0	1515	4	US-09-071-035-431	Sequence 431, App
4		15.2	76.0	1803	4	US-09-071-035-429	Sequence 429, App
5		15.2	76.0	2481	4	US-09-134-000C-3193	Sequence 3193, Ap
6		15.2	76.0	2611	4	US-09-620-312D-925	Sequence 925, App
c 7		15.2	76.0	3614	4	US-09-221-013A-9	Sequence 9, Appli
c 8		15.2	76.0	48974	3	US-08-920-422-17	Sequence 17, Appl
c 9		14.8	74.0	514	4	US-09-621-976-14354	Sequence 14354, A
c 10		14.8	74.0	2068	2	US-08-466-589-1	Sequence 1, Appli
c 11		14.8	74.0	2068	2	US-08-700-636-1	Sequence 1, Appli
c 12		14.8	74.0	2068	3	US-08-467-574-1	Sequence 1, Appli
c 13		14.8	74.0	2068	4	US-09-217-345-1	Sequence 1, Appli
c 14		14.8	74.0	2068	4	US-09-892-985-1	Sequence 1, Appli
c 15		14.8	74.0	2277	1	US-08-496-855A-1	Sequence 1, Appli
c 16		14.8	74.0	2277	4	US-08-487-596-1	Sequence 1, Appli
c 17		14.8	74.0	2352	2	US-08-889-909A-21	Sequence 21, Appl
c 18		14.8	74.0	2352	4	US-09-156-163A-21	Sequence 21, Appl
c 19		14.8	74.0	2352	4	US-09-982-308B-21	Sequence 21, Appl
c 20		14.8	74.0	2430	1	US-08-062-368-1	Sequence 1, Appli
c 21		14.8	74.0	2664	4	US-08-660-451A-1	Sequence 1, Appli
c 22		14.8	74.0	6268	4	US-09-566-921-57	Sequence 57, Appl
23		14.8	74.0	18994	1	US-08-459-586-4	Sequence 4, Appli
24		14.8	74.0	18994	2	US-08-282-696-4	Sequence 4, Appli
25		14.4	72.0	34063	4	US-09-453-702B-96	Sequence 96, Appl
c 26		14.4	72.0	36519	3	US-08-923-137-2	Sequence 2, Appli
c 27		14.2	71.0	998	4	US-09-671-317-191	Sequence 191, App
28		14.2	71.0	1229	4	US-09-404-641-80	Sequence 80, Appl
29		14.2	71.0	1298	4	US-09-404-641-69	Sequence 69, Appl
30		14.2	71.0	1398	4	US-09-328-352-2591	Sequence 2591, Ap
c 31		14.2	71.0	1572	4	US-09-620-312D-886	Sequence 886, App
32		14.2	71.0	1735	4	US-09-404-641-84	Sequence 84, Appl
33		14.2	71.0	1773	4	US-09-134-000C-1350	Sequence 1350, Ap
34		14.2	71.0	1799	4	US-09-732-234-5	Sequence 5, Appli
35		14.2	71.0	1799	4	US-09-784-859-5	Sequence 5, Appli
c 36		14.2	71.0	3588	4	US-09-620-312D-863	Sequence 863, App
c 37		14.2	71.0	41171	4	US-08-311-731A-122	Sequence 122, App
38		14.2	71.0	1664976	4	US-08-916-421B-1	Sequence 1, Appli
c 39		14.2	71.0	1664976	4	US-08-916-421B-1	Sequence 1, Appli
c 40		14	70.0	1326	4	US-09-328-352-3788	Sequence 3788, Ap
41		13.8	69.0	159	3	US-08-651-136C-47	Sequence 47, Appl
42		13.8	69.0	159	4	US-09-229-911A-47	Sequence 47, Appl
43		13.8	69.0	180	3	US-08-651-136C-69	Sequence 69, Appl
44		13.8	69.0	180	4	US-09-229-911A-69	Sequence 69, Appl
c 45		13.8	69.0	291	4	US-09-621-976-17092	Sequence 17092, A

ALIGNMENTS

RESULT 1
US-09-816-095-3
; Sequence 3, Application US/09816095
; Patent No. 6664084
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001147
; CURRENT APPLICATION NUMBER: US/09/816,095
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 99916
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(99916)
; OTHER INFORMATION: n = A,T,C or G
US-09-816-095-3

Query Match 79.0%; Score 15.8; DB 4; Length 99916;
Best Local Similarity 89.5%; Pred. No. 86;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          2 TGGACAGGTTAGGGCTTTG 20
              ||||||| ||||| |||||
Db       5532 TGGACAGATTAGGCCTTTG 5550
```

RESULT 2
US-09-489-039A-4299/c
; Sequence 4299, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 4299
; LENGTH: 303
; TYPE: DNA

```
;   ORGANISM: Klebsiella pneumoniae
US-09-489-039A-4299

  Query Match             76.0%;  Score 15.2;  DB 4;  Length 303;
  Best Local Similarity   85.0%;  Pred. No. 69;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CTGGACAGGTTAGGGCTTTG 20
              ||| |||||| || ||||||
Db        250 CTGCACAGGTCAGAGCTTTG 231


RESULT 3
US-09-071-035-431
; Sequence 431, Application US/09071035
; Patent No. 6448043
;  GENERAL INFORMATION:
;    APPLICANT:  Gil H. Choi
;    TITLE OF INVENTION:  Enterococcus faecalis Polynucleotides and Polypeptides
;    NUMBER OF SEQUENCES:  496
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/071,035
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  A. Anders Brookes
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PB369P2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  431:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1515 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-09-071-035-431

  Query Match             76.0%;  Score 15.2;  DB 4;  Length 1515;
  Best Local Similarity   85.0%;  Pred. No. 89;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CTGGACAGGTTAGGGCTTTG 20
              ||||||||||  || |||||
Db        430 CTGGACAGGTGCGGTCTTTG 449


RESULT 4
US-09-071-035-429
; Sequence 429, Application US/09071035
; Patent No. 6448043
;  GENERAL INFORMATION:
;    APPLICANT:  Gil H. Choi
;    TITLE OF INVENTION:  Enterococcus faecalis Polynucleotides and Polypeptides
;    NUMBER OF SEQUENCES:  496
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/071,035
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  A. Anders Brookes
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PB369P2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  429:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1803 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-09-071-035-429

  Query Match             76.0%;  Score 15.2;  DB 4;  Length 1803;
  Best Local Similarity   85.0%;  Pred. No. 92;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CTGGACAGGTTAGGGCTTTG 20
              ||||||||||  || |||||
Db        690 CTGGACAGGTGCGGTCTTTG 709


RESULT 5
US-09-134-000C-3193
; Sequence 3193, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
;  APPLICANT: Lynn Doucette-Stamm et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;  TITLE OF INVENTION:  ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 032796-032
;  CURRENT APPLICATION NUMBER: US/09/134,000C
;  CURRENT FILING DATE:  1998-08-13
;  PRIOR APPLICATION NUMBER: US 60/055,778
;  PRIOR FILING DATE: 1997-08-15
;  NUMBER OF SEQ ID NOS: 6812
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3193
;   LENGTH: 2481
;   TYPE: DNA
;   ORGANISM: Enterococcus faecalis
US-09-134-000C-3193

  Query Match             76.0%;  Score 15.2;  DB 4;  Length 2481;
  Best Local Similarity   85.0%;  Pred. No. 97;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CTGGACAGGTTAGGGCTTTG 20
              ||||||||||  || |||||
Db        687 CTGGACAGGTGCGGTCTTTG 706


RESULT 6
US-09-620-312D-925
```

```
; Sequence 925, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
;  APPLICANT: Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Asundi, Vinod
;  APPLICANT:  Zhang, Jie
;  APPLICANT:  Ren, Feiyan
;  APPLICANT:  Chen, Rui-hong
;  APPLICANT:  Zhao, Qing A.
;  APPLICANT:  Wehrman, Tom
;  APPLICANT:  Xue, Aidong J.
;  APPLICANT:  Yang, Yonghong
;  APPLICANT:  Wang, Jian-Rui
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Ma, Yunqing
;  APPLICANT:  Wang, Dunrui
;  APPLICANT:  Wang, Zhiwei
;  APPLICANT:  John Tillinghast
;  APPLICANT:  Drmanac, Radoje T.
;  TITLE OF INVENTION: No. 6569662el Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides
;  FILE REFERENCE: 784CIP2B
;  CURRENT APPLICATION NUMBER: US/09/620,312D
;  CURRENT FILING DATE:  2000-07-19
;  PRIOR APPLICATION NUMBER: 09/552,317
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: 09/488,725
;  PRIOR FILING DATE: 2000-01-21
;  NUMBER OF SEQ ID NOS: 1105
;  SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 925
;   LENGTH: 2611
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (290)..(1885)
US-09-620-312D-925

  Query Match             76.0%;  Score 15.2;  DB 4;  Length 2611;
  Best Local Similarity   85.0%;  Pred. No. 98;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CTGGACAGGTTAGGGCTTTG 20
              ||| |||| ||||| |||||
Db        789 CTGAACAGATTAGGTCTTTG 808


RESULT 7
US-09-221-013A-9/c
; Sequence 9, Application US/09221013A
; Patent No. 6495740
; GENERAL INFORMATION:
;  APPLICANT:  Arioli, Antonio
;  APPLICANT:  Williamson, Richard E.
;  APPLICANT:  Betzner, Andreas S.
;  APPLICANT:  Peng, Liangcai
;  TITLE OF INVENTION:  Manipulation of cellulose and/or Beta-1,4-glucan
;  FILE REFERENCE:  96-98
;  CURRENT APPLICATION NUMBER: US/09/221,013A
;  CURRENT FILING DATE:  1998-12-23
;  PRIOR APPLICATION NUMBER:  PCT/AU97/00402
;  PRIOR FILING DATE:  1997-06-24
;  PRIOR APPLICATION NUMBER:  AU PO0699
;  PRIOR FILING DATE:  1996-06-27
;  NUMBER OF SEQ ID NOS:  37
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 3614
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (217)..(3411)
US-09-221-013A-9

  Query Match             76.0%;  Score 15.2;  DB 4;  Length 3614;
  Best Local Similarity   85.0%;  Pred. No. 1e+02;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CTGGACAGGTTAGGGCTTTG 20
              | |||||  ||||||||||
Db       1600 CAGGACATTTTAGGGCTTTG 1581


RESULT 8
US-08-920-422-17/c
; Sequence 17, Application US/08920422A
; Patent No. 6255473
; GENERAL INFORMATION:
;  APPLICANT: Vitek, Michael P.
;  APPLICANT: Mitsuda, No. 6255473iaki
;  APPLICANT: Roses, Allen D.
;  TITLE OF INVENTION: Presenilin-1 Gene Promoter
;  FILE REFERENCE: VITEKPRESENILIN
;  CURRENT APPLICATION NUMBER: US/08/920,422A
;  CURRENT FILING DATE: 1997-08-29
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
;   LENGTH: 48974
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-08-920-422-17

  Query Match             76.0%;  Score 15.2;  DB 3;  Length 48974;
  Best Local Similarity   85.0%;  Pred. No. 1.6e+02;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CTGGACAGGTTAGGGCTTTG 20
              |||| |||| ||||||| ||
Db      33769 CTGGCCAGGATAGGGCTGTG 33750


RESULT 9
US-09-621-976-14354/c
; Sequence 14354, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Jobert, S.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: ESTs and Encoded Human Proteins.
;  FILE REFERENCE: GENSET.054PR2
;  CURRENT APPLICATION NUMBER: US/09/621,976
;  CURRENT FILING DATE:  2000-07-21
;  NUMBER OF SEQ ID NOS: 19335
;  SOFTWARE: Patent.pm
; SEQ ID NO 14354
;   LENGTH: 514
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 254
;   OTHER INFORMATION: n=a, g, c or t
US-09-621-976-14354

  Query Match             74.0%;  Score 14.8;  DB 4;  Length 514;
  Best Local Similarity   88.9%;  Pred. No. 1.2e+02;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 GGACAGGTTAGGGCTTTG 20
              ||||||| |||||||| |
Db        311 GGACAGGGTAGGGCTTGG 294
```

```
RESULT 10
US-08-466-589-1/c
; Sequence 1, Application US/08466589
; Patent No. 5837489
;  GENERAL INFORMATION:
;    APPLICANT:  Elliot, Kathryn J.
;    APPLICANT:  Ellis, Steven B.
;    APPLICANT:  Harpold, Michael M.
;    TITLE OF INVENTION:  HUMAN NEURONAL NICOTINIC ACETYLCHOLINE
;    TITLE OF INVENTION:  RECEPTOR COMPOSITIONS AND METHODS EMPLOYING SAME
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClaim
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/466,589
;      FILING DATE:  June 5, 1995
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/028,031
;      FILING DATE:  March 8, 1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L
;      REGISTRATION NUMBER:  33,779
;      REFERENCE/DOCKET NUMBER:  6362-9950
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-238-0999
;      TELEFAX:  619-238-0062
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2068 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  166..1752
US-08-466-589-1

  Query Match             74.0%;  Score 14.8;  DB 2;  Length 2068;
  Best Local Similarity   88.9%;  Pred. No. 1.5e+02;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 GGACAGGTTAGGGCTTTG 20
              || ||||| |||||||||
Db        144 GGTCAGGTCAGGGCTTTG 127


RESULT 11
US-08-700-636-1/c
; Sequence 1, Application US/08700636
; Patent No. 5910582
;  GENERAL INFORMATION:
;    APPLICANT:  Elliot, Kathryn J.
;    APPLICANT:  Ellis, Steven B.
;    APPLICANT:  Harpold, Michael M.
;    TITLE OF INVENTION:  HUMAN NEURONAL NICOTINIC ACETYLCHOLINE
;    TITLE OF INVENTION:  RECEPTOR COMPOSITIONS AND METHODS EMPLOYING SAME
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pretty, Schroeder, Brueggemann & Clark
;      STREET:  444 South Flower Street, Suite 2000
;      CITY:  Los Angeles
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/700,636
;      FILING DATE:  16-JUL-1996
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/028,031
;      FILING DATE:  08-MAR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Reiter, Stephen E.
;      REGISTRATION NUMBER:  31,192
;      REFERENCE/DOCKET NUMBER:  P41 9368
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-546-4737
;      TELEFAX:  619-546-9392
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2068 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  166..1752
US-08-700-636-1

  Query Match             74.0%;  Score 14.8;  DB 2;  Length 2068;
  Best Local Similarity   88.9%;  Pred. No. 1.5e+02;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 GGACAGGTTAGGGCTTTG 20
              || ||||| |||||||||
Db        144 GGTCAGGTCAGGGCTTTG 127


RESULT 12
US-08-467-574-1/c
; Sequence 1, Application US/08467574
; Patent No. 6022704
;  GENERAL INFORMATION:
;    APPLICANT:  Elliot, Kathryn J.
;    APPLICANT:  Ellis, Steven B.
;    APPLICANT:  Harpold, Michael M.
;    TITLE OF INVENTION:  HUMAN NEURONAL NICOTINIC ACETYLCHOLINE
;    TITLE OF INVENTION:  RECEPTOR COMPOSITIONS AND METHODS EMPLOYING SAME
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClaim
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/467,574
;      FILING DATE:  June 5, 1995
;      CLASSIFICATION:  536
```

```
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/028,031
;       FILING DATE:  March 8, 1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Seidman, Stephanie L
;       REGISTRATION NUMBER:  33,779
;       REFERENCE/DOCKET NUMBER:  6362-9949
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  619-238-0999
;       TELEFAX:  619-238-0062
;       TELEX:
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  2068 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  both
;       TOPOLOGY:  both
;     MOLECULE TYPE:  cDNA
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  166..1752
US-08-467-574-1

  Query Match             74.0%;  Score 14.8;  DB 3;  Length 2068;
  Best Local Similarity   88.9%;  Pred. No. 1.5e+02;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 GGACAGGTTAGGGCTTTG 20
              || ||||| |||||||||
Db        144 GGTCAGGTCAGGGCTTTG 127


RESULT 13
US-09-217-345-1/c
; Sequence 1, Application US/09217345
; Patent No. 6303753
;  GENERAL INFORMATION:
;    APPLICANT:  Elliot, Kathryn J.
;    APPLICANT:  Ellis, Steven B.
;    APPLICANT:  Harpold, Michael M.
;    TITLE OF INVENTION:  HUMAN NEURONAL NICOTINIC ACETYLCHOLINE
;    TITLE OF INVENTION:  RECEPTOR COMPOSITIONS AND METHODS EMPLOYING SAME
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Heller Ehrman White & McAuliffe
;      STREET:  4250 Executive Square, 7th Floor
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/217,345
;      FILING DATE:  21-DEC-98
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/467,574
;      FILING DATE:  05-JUN-95
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/466,589,
;      FILING DATE:  05-JUN-95
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/028,031
;      FILING DATE:  08-MAR-93
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L
;      REGISTRATION NUMBER:  33,779
;      REFERENCE/DOCKET NUMBER:  24735-9949B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-450-8400
;      TELEFAX:  619-587-5360
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2068 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  166..1752
US-09-217-345-1

  Query Match             74.0%;  Score 14.8;  DB 4;  Length 2068;
  Best Local Similarity   88.9%;  Pred. No. 1.5e+02;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 GGACAGGTTAGGGCTTTG 20
              || ||||| |||||||||
Db        144 GGTCAGGTCAGGGCTTTG 127


RESULT 14
US-09-892-985-1/c
; Sequence 1, Application US/09892985
; Patent No. 6664375
;  GENERAL INFORMATION:
;       APPLICANT: Elliot, Kathryn J.
;                  Ellis, Steven B.
;                  Harpold, Michael M.
;       TITLE OF INVENTION: HUMAN NEURONAL NICOTINIC ACETYLCHOLINE
;                           RECEPTOR COMPOSITIONS AND METHODS EMPLOYING SAME
;       NUMBER OF SEQUENCES: 12
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Heller Ehrman White & McAuliffe
;            STREET: 4250 Executive Square, 7th Floor
;            CITY: La Jolla
;            STATE: CA
;            COUNTRY: USA
;            ZIP: 92037
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette
;            COMPUTER: IBM Compatible
;            OPERATING SYSTEM: DOS
;            SOFTWARE: FastSEQ Version 1.5
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/892,985
;            FILING DATE: 27-Jun-2001
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/09/217,345
;            FILING DATE: 21-DEC-98
;            APPLICATION NUMBER: US 08/467,574
;            FILING DATE: 05-JUN-95
;            APPLICATION NUMBER: US 08/466,589,
;            FILING DATE: 05-JUN-95
;            APPLICATION NUMBER: US 08/028,031
;            FILING DATE: 08-MAR-93
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Seidman, Stephanie L
;            REGISTRATION NUMBER: 33,779
;            REFERENCE/DOCKET NUMBER: 24735-9949B
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 619-450-8400
;            TELEFAX: 619-587-5360
;            TELEX: <Unknown>
;  INFORMATION FOR SEQ ID NO: 1:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 2068 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: both
;            TOPOLOGY: both
;       MOLECULE TYPE: cDNA
```

```
;        FEATURE:
;             NAME/KEY:  CDS
;             LOCATION:  166..1752
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-892-985-1

  Query Match             74.0%;  Score 14.8;  DB 4;  Length 2068;
  Best Local Similarity   88.9%;  Pred. No. 1.5e+02;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 GGACAGGTTAGGGCTTTG 20
              || ||||| |||||||||
Db        144 GGTCAGGTCAGGGCTTTG 127


RESULT 15
US-08-496-855A-1/c
; Sequence 1, Application US/08496855A
; Patent No. 5801232
;  GENERAL INFORMATION:
;    APPLICANT:  Elliot, Kathryn J.
;    APPLICANT:  Ellis, Steven B.
;    APPLICANT:  Harpold, Michael M.
;    TITLE OF INVENTION:  HUMAN NEURONAL NICOTINIC ACETYLCHOLINE
;    TITLE OF INVENTION:  RECEPTOR COMPOSITIONS AND METHODS EMPLOYING SAME
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClain
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  CA
;      COUNTRY:  U.S.A.
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/496,855A
;      FILING DATE:  20-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/149,503
;      FILING DATE:  08-NOV-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/028,031
;      FILING DATE:  08-MAR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie
;      REGISTRATION NUMBER:  33,779
;      REFERENCE/DOCKET NUMBER:  6362-9369B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-238-0999
;      TELEFAX:  619-238-0062
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2277 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  166..1755
US-08-496-855A-1

  Query Match             74.0%;  Score 14.8;  DB 1;  Length 2277;
  Best Local Similarity   88.9%;  Pred. No. 1.5e+02;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 GGACAGGTTAGGGCTTTG 20
              || ||||| |||||||||
Db        144 GGTCAGGTCAGGGCTTTG 127


Search completed: February 20, 2004, 01:45:09
Job time : 6.2454 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 23:23:34 ; Search time 48.0735 Seconds
(without alignments)
1456.787 Million cell updates/sec

Title: US-09-939-853A-140
Perfect score: 20
Sequence: 1 ctggacaggttagggctttg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2308684 seqs, 1750822206 residues

Total number of hits satisfying chosen parameters: 4617368

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	20	100.0	444	9	US-09-867-550-951	Sequence 951, App
c	2	20	100.0	763	9	US-09-867-550-953	Sequence 953, App
c	3	20	100.0	864	10	US-09-814-353-21302	Sequence 21302, A
	4	17.4	87.0	422	15	US-10-242-535A-25371	Sequence 25371, A
	5	16.8	84.0	665	15	US-10-027-632-133814	Sequence 133814,
c	6	16.8	84.0	2305	15	US-10-094-749-795	Sequence 795, App
c	7	16.8	84.0	2424	15	US-10-027-632-103042	Sequence 103042,
c	8	16.8	84.0	2424	15	US-10-027-632-103043	Sequence 103043,
c	9	16.8	84.0	3559	15	US-10-108-260A-602	Sequence 602, App
	10	15.8	79.0	403	10	US-09-918-995-35904	Sequence 35904, A
	11	15.8	79.0	467	10	US-09-918-995-27228	Sequence 27228, A
	12	15.8	79.0	570	15	US-10-027-632-137211	Sequence 137211,
	13	15.8	79.0	663	15	US-10-027-632-208024	Sequence 208024,
	14	15.8	79.0	1152	10	US-09-882-227-295	Sequence 295, App
	15	15.8	79.0	1744	12	US-10-424-599-121358	Sequence 121358,
	16	15.8	79.0	1824	12	US-10-424-599-73670	Sequence 73670, A
	17	15.8	79.0	4170	10	US-09-919-039-221	Sequence 221, App
	18	15.8	79.0	4170	14	US-10-168-425-23	Sequence 23, Appl
	19	15.8	79.0	4359	15	US-10-191-803-341	Sequence 341, App
	20	15.8	79.0	24023	13	US-10-094-679-1	Sequence 1, Appli
	21	15.8	79.0	99916	9	US-09-816-095-3	Sequence 3, Appli
	22	15.4	77.0	673	9	US-09-917-800A-1299	Sequence 1299, Ap
	23	15.4	77.0	761	15	US-10-027-632-144930	Sequence 144930,
	24	15.4	77.0	786	15	US-10-027-632-169995	Sequence 169995,
	25	15.4	77.0	786	15	US-10-027-632-169996	Sequence 169996,
c	26	15.4	77.0	978	15	US-10-027-632-121171	Sequence 121171,
c	27	15.4	77.0	978	15	US-10-027-632-121172	Sequence 121172,
	28	15.4	77.0	2067	14	US-10-101-510-405	Sequence 405, App
	29	15.4	77.0	3685	14	US-10-225-567A-610	Sequence 610, App
c	30	15.4	77.0	4112	10	US-09-930-213-301	Sequence 301, App
	31	15.4	77.0	136726	15	US-10-085-117-244	Sequence 244, App
c	32	15.2	76.0	309	9	US-09-836-607-7	Sequence 7, Appli
c	33	15.2	76.0	309	10	US-09-421-112-7	Sequence 7, Appli
c	34	15.2	76.0	320	14	US-10-106-698-3422	Sequence 3422, Ap
c	35	15.2	76.0	430	9	US-09-764-869-408	Sequence 408, App
c	36	15.2	76.0	430	14	US-10-091-504-408	Sequence 408, App
c	37	15.2	76.0	430	15	US-10-227-577-408	Sequence 408, App
c	38	15.2	76.0	468	9	US-09-836-607-34	Sequence 34, Appl
c	39	15.2	76.0	468	10	US-09-421-112-34	Sequence 34, Appl
c	40	15.2	76.0	472	10	US-09-918-995-20214	Sequence 20214, A
c	41	15.2	76.0	475	10	US-09-918-995-1752	Sequence 1752, Ap
c	42	15.2	76.0	478	10	US-09-918-995-32274	Sequence 32274, A
	43	15.2	76.0	547	15	US-10-027-632-36047	Sequence 36047, A
c	44	15.2	76.0	547	15	US-10-027-632-61127	Sequence 61127, A
c	45	15.2	76.0	547	15	US-10-027-632-303041	Sequence 303041,

ALIGNMENTS

RESULT 1
US-09-867-550-951/c
; Sequence 951, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells and I
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 951
; LENGTH: 444
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-867-550-951

Query Match 100.0%; Score 20; DB 9; Length 444;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CTGGACAGGTTAGGGCTTTG 20
              ||||||||||||||||||||
Db         60 CTGGACAGGTTAGGGCTTTG 41
```

RESULT 2
US-09-867-550-953/c
; Sequence 953, Application US/09867550
; Patent No. US20020082206A1

```
; GENERAL INFORMATION:
;  APPLICANT: Leach, Martin D.
;  APPLICANT:  Mehraban, Fuad,
;  APPLICANT:  Conley, Pamela
;  APPLICANT:  Law, Debbie
;  APPLICANT:  Topper, James
;  TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells and
;  TITLE OF INVENTION:  Thereby
;  FILE REFERENCE: 21402-013 (Cura-313)
;  CURRENT APPLICATION NUMBER: US/09/867,550
;  CURRENT FILING DATE:  2001-09-20
;  PRIOR APPLICATION NUMBER: USSN 60/208,427
;  PRIOR FILING DATE: 2000-05-30
;  NUMBER OF SEQ ID NOS: 2125
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 953
;   LENGTH: 763
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-867-550-953

  Query Match             100.0%;  Score 20;  DB 9;  Length 763;
  Best Local Similarity   100.0%;  Pred. No. 1.1;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGGACAGGTTAGGGCTTTG 20
              ||||||||||||||||||||
Db        189 CTGGACAGGTTAGGGCTTTG 170


RESULT 3
US-09-814-353-21302/c
; Sequence 21302, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
;  APPLICANT: Lee, John
;  APPLICANT:  Thompson, Pamela
;  APPLICANT:  Lillie, James
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
;  TITLE OF INVENTION:  IDENTIFICATION, ASSESSMENT, PREVENTION, AND
;  TITLE OF INVENTION:  THERAPY OF OVARIAN CANCER
;  FILE REFERENCE: MRI-006B
;  CURRENT APPLICATION NUMBER: US/09/814,353
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: US 60/191,031
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: US 60/207,124
;  PRIOR FILING DATE: 2000-05-25
;  PRIOR APPLICATION NUMBER: US 60/211,940
;  PRIOR FILING DATE: 2000-06-15
;  PRIOR APPLICATION NUMBER: US 60/216,820
;  PRIOR FILING DATE: 2000-07-07
;  PRIOR APPLICATION NUMBER: US 60/220,661
;  PRIOR FILING DATE: 2000-07-25
;  PRIOR APPLICATION NUMBER: US 60/257,672
;  PRIOR FILING DATE: 2000-12-21
;  NUMBER OF SEQ ID NOS: 22037
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 21302
;   LENGTH: 864
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 1, 2, 3, 32, 862, 863, 864
;   OTHER INFORMATION: n = A,T,C or G
US-09-814-353-21302

  Query Match             100.0%;  Score 20;  DB 10;  Length 864;
  Best Local Similarity   100.0%;  Pred. No. 1.1;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGGACAGGTTAGGGCTTTG 20
              ||||||||||||||||||||
Db        353 CTGGACAGGTTAGGGCTTTG 334


RESULT 4
US-10-242-535A-25371
; Sequence 25371, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
;  APPLICANT: ChondroGene Inc.
;  APPLICANT:  Liew, C.C.
;  TITLE OF INVENTION: Compositions and Methods Relatiing to Osteoarthritis
;  FILE REFERENCE: 4231/2005
;  CURRENT APPLICATION NUMBER: US/10/242,535A
;  CURRENT FILING DATE:  2002-09-12
;  PRIOR APPLICATION NUMBER: US 10/085,783
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/305,340
;  PRIOR FILING DATE: 2001-07-13
;  PRIOR APPLICATION NUMBER: US 60/275,017
;  PRIOR FILING DATE: 2001-03-12
;  PRIOR APPLICATION NUMBER: US 60/271,955
;  PRIOR FILING DATE: 2001-02-28
;  NUMBER OF SEQ ID NOS: 58994
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 25371
;   LENGTH: 422
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (5)..(5)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (7)..(7)
;   OTHER INFORMATION: n is a, c, g, or t
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (406)..(406)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-25371

  Query Match             87.0%;  Score 17.4;  DB 15;  Length 422;
  Best Local Similarity   94.7%;  Pred. No. 24;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 CTGGACAGGTTAGGGCTTT 19
              ||||| ||||||||||||||
Db         61 CTGGAGAGGTTAGGGCTTT 79


RESULT 5
US-10-027-632-133814
; Sequence 133814, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
```

```
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 133814
;   LENGTH: 665
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-133814

  Query Match             84.0%;  Score 16.8;  DB 15;  Length 665;
  Best Local Similarity   90.0%;  Pred. No. 51;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 CTGGACAGGTTAGGGCTTTG 20
              |||||||||  |||||||  ||
Db        344 CTGGACAGGATAGGGCTGTG 363


RESULT 6
US-10-094-749-795/c
; Sequence 795, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
;  APPLICANT: ISOGAI, TAKAO
;  APPLICANT:  SUGIYAMA, TOMOYASU
;  APPLICANT:  OTSUKI, TETSUJI
;  APPLICANT:  WAKAMATSU, AI
;  APPLICANT:  SATO, HIROYUKI
;  APPLICANT:  ISHII, SHIZUKO
;  APPLICANT:  YAMAMOTO, JUN-ICHI
;  APPLICANT:  ISONO, YUUKO
;  APPLICANT:  HIO, YURI
;  APPLICANT:  OTSUKA, KAORU
;  APPLICANT:  NAGAI, KEIICHI
;  APPLICANT:  IRIE, RYOTARO
;  APPLICANT:  TAMECHIKA, ICHIRO
;  APPLICANT:  SEKI, NAOHIKO
;  APPLICANT:  YOSHIKAWA, TSUTOMU
;  APPLICANT:  OTSUKA, MOTOYUKI
;  APPLICANT:  NAGAHARI, KENJI
;  APPLICANT:  MASUHO, YASUHIKO
;  TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
;  FILE REFERENCE: 084335/0160
;  CURRENT APPLICATION NUMBER: US/10/094,749
;  CURRENT FILING DATE:  2002-03-12
;  PRIOR APPLICATION NUMBER: 60/350,435
;  PRIOR FILING DATE: 2002-01-24
;  PRIOR APPLICATION NUMBER: JP 2001-328381
;  PRIOR FILING DATE: 2001-09-14
;  NUMBER OF SEQ ID NOS: 3381
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 795
;   LENGTH: 2305
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-094-749-795

  Query Match             84.0%;  Score 16.8;  DB 15;  Length 2305;
  Best Local Similarity   90.0%;  Pred. No. 53;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 CTGGACAGGTTAGGGCTTTG 20
              |||||||| ||||||||  ||
Db       2118 CTGGACAGTTTAGGGCTGTG 2099


RESULT 7
US-10-027-632-103042/c
; Sequence 103042, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 103042
;   LENGTH: 2424
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-103042

  Query Match             84.0%;  Score 16.8;  DB 15;  Length 2424;
  Best Local Similarity   90.0%;  Pred. No. 53;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 CTGGACAGGTTAGGGCTTTG 20
              |||||||| ||||||||  ||
Db       1453 CTGGACAGTTTAGGGCTGTG 1434


RESULT 8
US-10-027-632-103043/c
; Sequence 103043, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 103043
;   LENGTH: 2424
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-103043

  Query Match             84.0%;  Score 16.8;  DB 15;  Length 2424;
  Best Local Similarity   90.0%;  Pred. No. 53;
```

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 CTGGACAGGTTAGGGCTTTG 20
              |||||||| |||||||| ||
Db       1453 CTGGACAGTTTAGGGCTGTG 1434
```

RESULT 9
US-10-108-260A-602/c
; Sequence 602, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 602
; LENGTH: 3559
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-602

Query Match 84.0%; Score 16.8; DB 15; Length 3559;
Best Local Similarity 90.0%; Pred. No. 54;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 CTGGACAGGTTAGGGCTTTG 20
              ||| |||||| |||||||||
Db         32 CTGAACAGGTCAGGGCTTTG 13
```

RESULT 10
US-09-918-995-35904
; Sequence 35904, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 35904
; LENGTH: 403
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(403)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-35904

Query Match 79.0%; Score 15.8; DB 10; Length 403;
Best Local Similarity 89.5%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 CTGGACAGGTTAGGGCTTT 19
              ||||| |||||| ||||||
Db         48 CTGGAGAGGTTAAGGCTTT 66
```

RESULT 11
US-09-918-995-27228
; Sequence 27228, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 27228
; LENGTH: 467
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(467)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-27228

Query Match 79.0%; Score 15.8; DB 10; Length 467;
Best Local Similarity 89.5%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 CTGGACAGGTTAGGGCTTT 19
              ||||| |||||| ||||||
Db        152 CTGGAGAGGTTAAGGCTTT 170
```

RESULT 12
US-10-027-632-137211
; Sequence 137211, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 137211
; LENGTH: 570
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-137211

Query Match 79.0%; Score 15.8; DB 15; Length 570;
Best Local Similarity 89.5%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 CTGGACAGGTTAGGGCTTT 19
              | |||||| ||||||||||
Db        301 CAGGACAGATTAGGGCTTT 319
```

RESULT 13
US-10-027-632-208024

```
; Sequence 208024, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 208024
;   LENGTH: 663
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-208024

  Query Match             79.0%;  Score 15.8;  DB 15;  Length 663;
  Best Local Similarity   89.5%;  Pred. No. 1.7e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 CTGGACAGGTTAGGGCTTT 19
              ||||||| |||||| |||
Db        600 CTGGACAGCTTAGGGTTTT 618


RESULT 14
US-09-882-227-295
; Sequence 295, Application US/09882227
; Publication No. US20030158396A1
; GENERAL INFORMATION:
;  APPLICANT: Kleanthous, Harold
;  APPLICANT:  Al-Garawi, Amal
;  APPLICANT:  Miller, Charles
;  APPLICANT:  Tomb, Jean-Francois
;  APPLICANT:  Ooomen, Raymond P.
;  TITLE OF INVENTION: Identification of Polynucleotides
;  TITLE OF INVENTION:  Encoding No. US20030158396Aiel Helicobacter Polypeptides in the
;  TITLE OF INVENTION:  Genome
;  FILE REFERENCE: 06132/047002
;  CURRENT APPLICATION NUMBER: US/09/882,227
;  CURRENT FILING DATE:  2001-06-15
;  PRIOR APPLICATION NUMBER: US 08/902,615
;  PRIOR FILING DATE: 1997-07-29
;  NUMBER OF SEQ ID NOS: 638
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 295
;   LENGTH: 1152
;   TYPE: DNA
;   ORGANISM: Helicobacter pylori
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (52)...(1095)
US-09-882-227-295

  Query Match             79.0%;  Score 15.8;  DB 10;  Length 1152;
  Best Local Similarity   89.5%;  Pred. No. 1.7e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 CTGGACAGGTTAGGGCTTT 19
              || ||| |||||||||||||
Db        934 CTAGACGGGTTAGGGCTTT 952


RESULT 15
US-10-424-599-121358
; Sequence 121358, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
;  APPLICANT: La Rosa Thomas J
;  APPLICANT:  Kovalic David K
;  APPLICANT:  Zhou Yihua
;  APPLICANT:  Cao Yongwei
;  TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants and Uses Thereof for Plant Improvement
;  FILE REFERENCE: 38-21(53223)B
;  CURRENT APPLICATION NUMBER: US/10/424,599
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 121358
;   LENGTH: 1744
;   TYPE: DNA
;   ORGANISM: Glycine max
;   FEATURE:
;   OTHER INFORMATION: Clone ID: PAT_MRT3847_80594C.1
US-10-424-599-121358

  Query Match             79.0%;  Score 15.8;  DB 12;  Length 1744;
  Best Local Similarity   89.5%;  Pred. No. 1.7e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 CTGGACAGGTTAGGGCTTT 19
              ||||  |||||||||||||
Db        695 CTGGCGAGGTTAGGGCTTT 713


Search completed: February 20, 2004, 05:59:20
Job time : 61.0735 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 21:21:18 ; Search time 113.831 Seconds
(without alignments)
9899.970 Million cell updates/sec

Title: US-09-939-853A-141
Perfect score: 26
Sequence: 1 ccttctggaagtctgccagtgtcctt 26

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	26	100.0	26	6	AX443200	AX443200 Sequence
c 2	26	100.0	1183	6	AX443133	AX443133 Sequence
3	26	100.0	1183	6	AX443135	AX443135 Sequence
c 4	26	100.0	2538	9	BC042041	BC042041 Homo sapi
c 5	26	100.0	2567	6	AX452880	AX452880 Sequence
c 6	26	100.0	2788	6	AX780857	AX780857 Sequence
7	26	100.0	66741	9	HS460J8	AL031662 Human DNA
c 8	26	100.0	145833	2	AC026539	AC026539 Homo sapi
c 9	21.2	81.5	177004	2	AC129114	AC129114 Rattus no
c 10	21.2	81.5	179497	9	AC020636	AC020636 Homo sapi
c 11	21.2	81.5	259967	2	AC105625	AC105625 Rattus no
c 12	21.2	81.5	261492	2	AC095338	AC095338 Rattus no
13	20.8	80.0	76332	2	AC021312	AC021312 Homo sapi
c 14	20.8	80.0	184541	9	AC011405	AC011405 Homo sapi
c 15	20.4	78.5	120194	9	AL355474	AL355474 Human DNA
16	20.4	78.5	157289	2	AC021447	AC021447 Homo sapi
17	20.4	78.5	163132	9	AC025038	AC025038 Homo sapi
c 18	20.2	77.7	1617	8	AK071089	AK071089 Oryza sat
c 19	20.2	77.7	2808	8	AK072171	AK072171 Oryza sat
20	20.2	77.7	110000	2	AC106661_2	Continuation (3 of
21	20.2	77.7	110000	2	AC127890_2	Continuation (3 of
c 22	20.2	77.7	130616	2	AP003950	AP003950 Oryza sat
c 23	20.2	77.7	139848	2	AP004335	AP004335 Oryza sat
c 24	20.2	77.7	148542	8	AP005127	AP005127 Oryza sat
25	20.2	77.7	163096	9	CNS01DU1	AL133167 Human chr
c 26	20.2	77.7	166305	9	AC012447	AC012447 Homo sapi
c 27	20.2	77.7	167551	2	AC112819	AC112819 Rattus no
28	20.2	77.7	168135	2	AC120654	AC120654 Rattus no
c 29	20.2	77.7	225443	2	AC137194	AC137194 Rattus no
c 30	20.2	77.7	246040	2	AC103184	AC103184 Rattus no
31	20.2	77.7	263262	2	AC095135	AC095135 Rattus no
c 32	19.8	76.2	92587	9	AC096665	AC096665 Homo sapi
c 33	19.8	76.2	165349	2	AC093665	AC093665 Homo sapi
c 34	19.8	76.2	170240	9	AL160279	AL160279 Human DNA
c 35	19.8	76.2	194998	10	AL627075	AL627075 Mouse DNA
c 36	19.8	76.2	210164	2	AC010732	AC010732 Homo sapi
c 37	19.8	76.2	241802	2	AC147280	AC147280 Pan trogl
c 38	19.6	75.4	6159	4	SSC133742	AJ133742 Sus scrof
c 39	19.6	75.4	139958	2	AC141936	AC141936 Rattus no
40	19.6	75.4	174987	2	AC145778	AC145778 Sus scrof
c 41	19.6	75.4	183097	9	AC087358	AC087358 Homo sapi
42	19.6	75.4	183360	2	AC022525	AC022525 Homo sapi
43	19.6	75.4	184160	2	AC147367	AC147367 Mus muscu
c 44	19.6	75.4	209380	10	AC133186	AC133186 Mus muscu
45	19.6	75.4	255608	2	AC109405	AC109405 Rattus no

ALIGNMENTS

RESULT 1
AX443200
LOCUS AX443200 26 bp DNA linear PAT 02-JUL-2002
DEFINITION Sequence 141 from Patent WO0216599.
ACCESSION AX443200
VERSION AX443200.1 GI:21690595
KEYWORDS .
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1
AUTHORS Burgess,C.E., Conley,P.B., Grosse,W.M., Hart,M., Kekuda,R., Shimkets,R.A., Spytek,K.A., Szekeres,E.S., Tomlinson,J.E., Topper,J.N. and Yang,R.B.
TITLE Proteins and nucleic acids encoding same

```
  JOURNAL   Patent: WO 0216599-A 141 28-FEB-2002;
            Curagen Corporation (US) ; COR THERAPEUTICS, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .26
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="oligonucleotide primer"
ORIGIN

  Query Match             100.0%;  Score 26;  DB 6;  Length 26;
  Best Local Similarity   100.0%;  Pred. No. 0.56;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCTTCTGGAAGTCTGCCAGTGTCCTT 26
              ||||||||||||||||||||||||||
Db          1 CCTTCTGGAAGTCTGCCAGTGTCCTT 26


RESULT 2
AX443133/c
LOCUS       AX443133                1183 bp    DNA     linear   PAT 02-JUL-2002
DEFINITION  Sequence 74 from Patent WO0216599.
ACCESSION   AX443133
VERSION     AX443133.1  GI:21690555
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Burgess,C.E., Conley,P.B., Grosse,W.M., Hart,M., Kekuda,R.,
            Shimkets,R.A., Spytek,K.A., Szekeres,E.S., Tomlinson,J.E.,
            Topper,J.N. and Yang,R.B.
  TITLE     Proteins and nucleic acids encoding same
  JOURNAL   Patent: WO 0216599-A 74 28-FEB-2002;
            Curagen Corporation (US) ; COR THERAPEUTICS, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .1183
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             100.0%;  Score 26;  DB 6;  Length 1183;
  Best Local Similarity   100.0%;  Pred. No. 0.37;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCTTCTGGAAGTCTGCCAGTGTCCTT 26
              ||||||||||||||||||||||||||
Db        276 CCTTCTGGAAGTCTGCCAGTGTCCTT 251


RESULT 3
AX443135
LOCUS       AX443135                1183 bp    DNA     linear   PAT 02-JUL-2002
DEFINITION  Sequence 76 from Patent WO0216599.
ACCESSION   AX443135
VERSION     AX443135.1  GI:21690556
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Burgess,C.E., Conley,P.B., Grosse,W.M., Hart,M., Kekuda,R.,
            Shimkets,R.A., Spytek,K.A., Szekeres,E.S., Tomlinson,J.E.,
            Topper,J.N. and Yang,R.B.
  TITLE     Proteins and nucleic acids encoding same
  JOURNAL   Patent: WO 0216599-A 76 28-FEB-2002;
            Curagen Corporation (US) ; COR THERAPEUTICS, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .1183
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             100.0%;  Score 26;  DB 6;  Length 1183;
  Best Local Similarity   100.0%;  Pred. No. 0.37;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCTTCTGGAAGTCTGCCAGTGTCCTT 26
              ||||||||||||||||||||||||||
Db        908 CCTTCTGGAAGTCTGCCAGTGTCCTT 933


RESULT 4
BC042041/c
LOCUS       BC042041                2538 bp    mRNA    linear   PRI 07-OCT-2003
DEFINITION  Homo sapiens Src-like-adaptor 2, transcript variant 1, mRNA (cDNA
            clone MGC:49845 IMAGE:4429896), complete cds.
ACCESSION   BC042041
VERSION     BC042041.1  GI:27469842
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2538)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2  (bases 1 to 2538)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-DEC-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 88 Row: a Column: 20
```

```
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 28416422.
FEATURES              Location/Qualifiers
     source           1. .2538
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="MGC:49845 IMAGE:4429896"
                      /tissue_type="Prostate, adenocarcinoma."
                      /clone_lib="NIH_MGC_91"
                      /lab_host="DH10B"
                      /note="Vector: pCMV-SPORT6"
     gene             1. .2538
                      /gene="SLA2"
                      /note="synonyms: FLJ21992, SLAP-2, C20orf156, MGC49845,
                      SLAP2"
                      /db_xref="LocusID:84174"
                      /db_xref="MIM:606577"
     CDS              363. .1148
                      /codon_start=1
                      /product="Src-like-adaptor 2, isoform a"
                      /protein_id="AAH42041.1"
                      /db_xref="GI:27469843"
                      /db_xref="LocusID:84174"
                      /translation="MGSLPSRRKSLPSPSLSSSVQGQGPVTMEAERSKATAVALGSFP
                      AGGPAELSLRLGEPLTIVSEDGDWWTVLSEVSGREYNIPSVHVAKVSHGWLYEGLSRE
                      KAEELLLLPGNPGGAFLIRESQTRRGSYSLSVRLSRPASWDRIRHYRIHCLDNGWLYI
                      SPRLTFPSLQALVDHYSELADDICCLLKEPCVLQRAGPLPGKDIPLPVTVQRTPLNWK
                      ELDSSLLFSEAATGEESLLSEGLRESLSFYISLNDEAVSLDDA"
     misc_feature     468. .632
                      /note="SH3; Region: SH3 domain. SH3 (Src homology 3)
                      domains are often indicative of a protein involved in
                      signal transduction related to cytoskeletal organization.
                      First described in the Src cytoplasmic tyrosine kinase.
                      The structure is a partly opened beta barrel"
                      /db_xref="CDD:pfam00018"
     misc_feature     642. .890
                      /note="SH2; Region: SH2 domain"
                      /db_xref="CDD:pfam00017"
ORIGIN

  Query Match             100.0%;  Score 26;  DB 9;  Length 2538;
  Best Local Similarity   100.0%;  Pred. No. 0.34;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCTTCTGGAAGTCTGCCAGTGTCCTT 26
              ||||||||||||||||||||||||||
Db        241 CCTTCTGGAAGTCTGCCAGTGTCCTT 216


RESULT 5
AX452880/c
LOCUS       AX452880               2567 bp    DNA     linear   PAT 06-JUL-2002
DEFINITION  Sequence 1 from Patent WO0242457.
ACCESSION   AX452880
VERSION     AX452880.1  GI:21712520
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Chang,H., Yang,W.P., Wu,Y., Whitney,G.S., Perez-Villar,J.J. and
            Kanner,S.B.
  TITLE     Cloning and expression of human slap-2: a novel sh2/sh3
            domain-containing human slap homologue having immune cell-specific
            expression
  JOURNAL   Patent: WO 0242457-A 1 30-MAY-2002;
            Bristol-Myers Squibb Co. (US)
FEATURES              Location/Qualifiers
     source           1. .2567
                      /organism="Homo sapiens"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:9606"
ORIGIN

  Query Match             100.0%;  Score 26;  DB 6;  Length 2567;
  Best Local Similarity   100.0%;  Pred. No. 0.34;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCTTCTGGAAGTCTGCCAGTGTCCTT 26
              ||||||||||||||||||||||||||
Db        293 CCTTCTGGAAGTCTGCCAGTGTCCTT 268


RESULT 6
AX780857/c
LOCUS       AX780857               2788 bp    DNA     linear   PAT 14-JUL-2003
DEFINITION  Sequence 3014 from Patent WO03039443.
ACCESSION   AX780857
VERSION     AX780857.1  GI:32697851
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Haferlach,T., Schoch,C., Kern,W., Kohlmann,A., Schnittger,S.,
            Dugas,M., Eils,R., Brors,B. and Mergenthaler,S.
  TITLE     Novel genetic markers for leukemias
  JOURNAL   Patent: WO 03039443-A 3014 15-MAY-2003;
            Deutsches Krebsforschungszentrum (DE) ;
            Ludwig-Maximilian-Universitaet Muenchen (DE) ; Haferlach, Torsten,
            PD Dr. Dr. (DE) ; Schoch, Claudia (DE) ; Kern, Wolfgang (DE)
FEATURES              Location/Qualifiers
     source           1. .2788
                      /organism="Homo sapiens"
                      /mol_type="unassigned DNA"
                      /db_xref="taxon:9606"
ORIGIN

  Query Match             100.0%;  Score 26;  DB 6;  Length 2788;
  Best Local Similarity   100.0%;  Pred. No. 0.33;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCTTCTGGAAGTCTGCCAGTGTCCTT 26
              ||||||||||||||||||||||||||
Db        265 CCTTCTGGAAGTCTGCCAGTGTCCTT 240


RESULT 7
HS460J8
LOCUS       HS460J8                66741 bp    DNA     linear   PRI 23-JUL-2001
DEFINITION  Human DNA sequence from clone RP3-460J8 on chromosome
            20q11.21-11.23 Contains the 3' end of the gene for a novel protein
            similar to N-myc downstream regulated (NDRG1) the 5' end of a gene
            encoding a novel protein tyrosine kinase, ESTs, STSs and GSSs,
            complete sequence.
ACCESSION   AL031662
VERSION     AL031662.26  GI:9716901
KEYWORDS    HTG; NDRG1; SH2 domain.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 66741)
  AUTHORS   Skuce,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Aug 7, 2000 this sequence version replaced gi:6425549.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
```

corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/Chr20
IMPORTANT: This sequence is not the entire insert of clone RP3-460J8 It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.
The true right end of clone RP3-460J8 is at 66741 in this sequence.
The true left end of clone RP3-469A13 is at 41767 in this sequence.
The true right end of clone RP5-977B1 is at 100 in this sequence.
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP3-460J8 is from the library RPCI-3 constructed by the group of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pCYPAC2.

FEATURES Location/Qualifiers

source 1. .66741
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="q11.21-11.23"
/clone="RP3-460J8"
/clone_lib="RPCI-3"

misc_feature complement(50. .544)
/note="match: GSS: Em:B45150"

gene complement(240. .7128)
/gene="dJ977B1.1"

mRNA complement(join(<240. .339,6995. .7128))
/gene="dJ977B1.1"
/product="dJ460J8.2 (novel protein tyrosine kinase with Src homology 2 (SH2) domain)"
/note="match: cDNAs: Em:AK025645
match: ESTs: Em:BG178487"
/evidence=not_experimental

CDS complement(join(<240. .339,6995. .7085))
/gene="dJ977B1.1"
/note="Continues in Em:AL050318 as dJ977B1.2"
/codon_start=1
/evidence=not_experimental
/product="dJ460J8.2 (novel protein tyrosine kinase with Src homology 2 (SH2) domain)"
/protein_id="CAC44645.1"
/db_xref="GI:15020830"
/db_xref="GOA:Q9H6Q3"
/db_xref="SWISS-PROT:Q9H6Q3"
/translation="MGSLPSRRKSLPSPSLSSSVQGQGPVTMEAERSKATAVALGSFP AGGPAELSLRLGEPLTIVS"

repeat_region 536. .565
/note="15 copies 2 mer ag 93% conserved"

repeat_region 2375. .2402
/note="14 copies 2 mer ta 100% conserved"

misc_feature 2783. .3155
/note="match: GSS: Em:AQ807191"

misc_feature complement(6983. .7482)
/note="match: GSS: Em:AQ556467"

misc_feature complement(6985. .7492)
/note="match: GSS: Em:AQ556478"

misc_feature 7486. .7925
/note="match: STS: Em:HS427J1S"

repeat_region 10615. .10640
/note="13 copies 2 mer tt 92% conserved"

repeat_region 13593. .13961
/note="MER21B repeat: matches 422. .785 of consensus"

repeat_region 14248. .14297
/note="MER21B repeat: matches 374. .422 of consensus"

repeat_region 14594. .14614
/note="MER21B repeat: matches 355. .374 of consensus"

repeat_region 14746. .15078
/note="MER21B repeat: matches 1. .355 of consensus"

polyA_site complement(17516)

misc_feature 17519. .17686
/note="match: GSS: Em:AZ067993"

polyA_site complement(17521)

polyA_site complement(17522)

gene complement(17523. .54534)
/gene="dJ469A13.3"

mRNA complement(join(<17523. .19451,20551. .20589,22110. .22161, 22247. .22282,26087. .26134,30796. .30847,30986. .31051, 32039. .32142,37091. .37147,46566. .46652,48268. .48328, 50163. .50225,53242. .53362,54429. .>54534))
/gene="dJ469A13.3"
/product="dJ460J8.1 (continued from dJ469A13.3 in Em:AL132768)"
/note="match: cDNAs: Em:AB033922 Em:U52073 Em:D87953 Em:X92845 Em:AF004162 Em:U60593 Em:AF045564 Em:AB033921 Em:AF159092 Em:AF147402 Em:M59814
match: ESTs: Em:AA718726 Em:AA039000 Em:AW003952 Em:W89263 Em:AA325826 Em:AI230982 Em:AA162360 Em:AA445016 Em:AV002395 Em:T85147 Em:AI786615 Em:AA113437 Em:AI004026 Em:T88705 Em:AI786673 Em:AI181197 Em:AA764653 Em:AA316771 Em:AU035165 Em:AV002368"
/evidence=not_experimental

polyA_signal complement(17542. .17547)
/gene="dJ469A13.3"

misc_feature 17555. .17811
/note="match: STS: Em:G19945"

misc_feature join(19202. .19451,22110. .22282,26087. .26127,30796. .30847)
/note="match: STS: Em:G23762"

CDS complement(join(19270. .19451,20551. .20589,22110. .22161, 22247. .22282,26087. .26134,30796. .30847,30986. .31051, 32039. .32142,37091. .37147,46566. .46652,48268. .48328, 50163. .50225,53242. .53362,54429. .>54534))
/gene="dJ469A13.3"
/note="novel protein (FLJ13556) similar to N-myc downstream regulated (NDRG1)
match: proteins: Sw:Q62433 Sw:Q92597 Tr:Q9Z2L9 Sw:P97862"
/codon_start=1
/evidence=not_experimental
/product="dJ460J8.1 (continued from dJ469A13.3 in Em:AL132768)"
/protein_id="CAB65625.1"
/db_xref="GI:6687781"
/db_xref="GOA:Q9UGV2"
/db_xref="SWISS-PROT:Q9UGV2"
/translation="EHDIETTHGVVHVTIRGLPKGNRPVILTYHDIGLNHKSCFNAPF NFEDMQEITQHFAVCHVDAPGQQEGAPSFPTGYQYPTMDELAEMLPPVLTHLSLKSII GIGVGAGAYILSRFALNHPELVEGLVLINVDPCAKGWIDWAASKLSGLTTNVVDIILA HHFGQEELQANLDLIQTYRMHIAQDINQDNLQLFLNSYNGRRDLEIERPILGQNDNKS KTLKCSTLLVVGDNSPAVEAVVECNSRLNPINTTLLKMADCGGLPQVVQPGKLTEAFK YFLQGMGYIPYVQLSHLSTESVPSASMTRLARSRTHSTSSSLGSGESPFSRSVTSNQS DGTQESCESPDVLDRHQTMEVSC"

misc_feature complement(24246. .24780)
/gene="dJ469A13.3"
/note="match: GSS: Em:AQ592789"

misc_feature complement(27980. .28556)
/gene="dJ469A13.3"
/note="match: GSS: Em:AQ308867"

repeat_region 34313. .34420
/note="HY1 repeat: matches 1. .109 of consensus"

repeat_region 34505. .34560
/note="28 copies 2 mer ta 78% conserved"

repeat_region 34564. .34611

```
                     /note="24 copies 2 mer ta 79% conserved"
     repeat_region   42553. .43658
                     /note="MER52C repeat: matches 1. .1278 of consensus"
     misc_feature    47989. .48470
                     /note="match: GSS: Em:AQ672417"
     repeat_region   49391. .49480
                     /note="45 copies 2 mer ta 76% conserved"
     misc_feature    52922. .53120
                     /note="match: STS: Em:G04621"
     misc_feature    complement(53426. .54032)
                     /gene="dJ469A13.3"
                     /note="match: GSS: Em:AQ314824
                     match: STS: Em:G55853"
     misc_feature    54054. .54477
                     /note="match: GSS: Em:AQ310681"
     misc_feature    57509. .57838
                     /note="match: STS: Em:G07504"
     misc_feature    complement(57696. .58013)
                     /note="match: STS: Em:G07632 Em:G07634"
     repeat_region   58856. .59001
                     /note="73 copies 2 mer at 83% conserved"
     misc_feature    complement(58930. .59281)
                     /note="match: GSS: Em:AQ067563"
     repeat_region   59062. .59236
                     /note="TIGGER2 repeat: matches 2541. .2714 of consensus"
     repeat_region   59738. .60495
                     /note="TIGGER2 repeat: matches 1780. .2541 of consensus"
     repeat_region   60719. .61831
                     /note="MER11C repeat: matches 1. .1071 of consensus"
     repeat_region   62547. .63174

  Query Match             100.0%;  Score 26;  DB 9;  Length 66741;
  Best Local Similarity   100.0%;  Pred. No. 0.23;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCTTCTGGAAGTCTGCCAGTGTCCTT 26
              ||||||||||||||||||||||||||
Db      11665 CCTTCTGGAAGTCTGCCAGTGTCCTT 11690


RESULT 8
AC026539/c
LOCUS       AC026539              145833 bp    DNA     linear   HTG 27-APR-2000
DEFINITION  Homo sapiens chromosome 20 clone RP11-712N14 map 20, WORKING DRAFT
            SEQUENCE, 37 unordered pieces.
ACCESSION   AC026539
VERSION     AC026539.2  GI:7656813
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 145833)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 20, clone RP11-712N14
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 145833)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-MAR-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Apr 27, 2000 this sequence version replaced gi:7283243.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L7115
                Center clone name: 712_N_14
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 125577 bases at least Q40
                Consensus quality: 135703 bases at least Q30
                Consensus quality: 139593 bases at least Q20
                Insert size: 182000; agarose-fp
                Insert size: 142233; sum-of-contigs
                Quality coverage: 2.6 in Q20 bases; agarose-fp
                Quality coverage: 3.3 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 37 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1214: contig of 1214 bp in length
            *     1215     1314: gap of 100 bp
            *     1315     2673: contig of 1359 bp in length
            *     2674     2773: gap of 100 bp
            *     2774     4520: contig of 1747 bp in length
            *     4521     4620: gap of 100 bp
            *     4621     5961: contig of 1341 bp in length
            *     5962     6061: gap of 100 bp
            *     6062     7719: contig of 1658 bp in length
            *     7720     7819: gap of 100 bp
            *     7820     9799: contig of 1980 bp in length
            *     9800     9899: gap of 100 bp
            *     9900    11434: contig of 1535 bp in length
            *    11435    11534: gap of 100 bp
            *    11535    14382: contig of 2848 bp in length
            *    14383    14482: gap of 100 bp
            *    14483    17148: contig of 2666 bp in length
            *    17149    17248: gap of 100 bp
            *    17249    19239: contig of 1991 bp in length
            *    19240    19339: gap of 100 bp
            *    19340    21102: contig of 1763 bp in length
            *    21103    21202: gap of 100 bp
            *    21203    23371: contig of 2169 bp in length
            *    23372    23471: gap of 100 bp
            *    23472    25782: contig of 2311 bp in length
            *    25783    25882: gap of 100 bp
            *    25883    28824: contig of 2942 bp in length
            *    28825    28924: gap of 100 bp
            *    28925    31619: contig of 2695 bp in length
            *    31620    31719: gap of 100 bp
            *    31720    34679: contig of 2960 bp in length
            *    34680    34779: gap of 100 bp
            *    34780    37527: contig of 2748 bp in length
```

```
            *    37528    37627: gap of 100 bp
            *    37628    40146: contig of 2519 bp in length
            *    40147    40246: gap of 100 bp
            *    40247    43743: contig of 3497 bp in length
            *    43744    43843: gap of 100 bp
            *    43844    47133: contig of 3290 bp in length
            *    47134    47233: gap of 100 bp
            *    47234    51023: contig of 3790 bp in length
            *    51024    51123: gap of 100 bp
            *    51124    54935: contig of 3812 bp in length
            *    54936    55035: gap of 100 bp
            *    55036    59553: contig of 4518 bp in length
            *    59554    59653: gap of 100 bp
            *    59654    64428: contig of 4775 bp in length
            *    64429    64528: gap of 100 bp
            *    64529    69211: contig of 4683 bp in length
            *    69212    69311: gap of 100 bp
            *    69312    72901: contig of 3590 bp in length
            *    72902    73001: gap of 100 bp
            *    73002    76724: contig of 3723 bp in length
            *    76725    76824: gap of 100 bp
            *    76825    81179: contig of 4355 bp in length
            *    81180    81279: gap of 100 bp
            *    81280    87009: contig of 5730 bp in length
            *    87010    87109: gap of 100 bp
            *    87110    90855: contig of 3746 bp in length
            *    90856    90955: gap of 100 bp
            *    90956    96520: contig of 5565 bp in length
            *    96521    96620: gap of 100 bp
            *    96621   102321: contig of 5701 bp in length
            *   102322   102421: gap of 100 bp
            *   102422   108293: contig of 5872 bp in length
            *   108294   108393: gap of 100 bp
            *   108394   116689: contig of 8296 bp in length
            *   116690   116789: gap of 100 bp
            *   116790   125264: contig of 8475 bp in length
            *   125265   125364: gap of 100 bp
            *   125365   136354: contig of 10990 bp in length
            *   136355   136454: gap of 100 bp
            *   136455   145833: contig of 9379 bp in length.
FEATURES             Location/Qualifiers
     source          1. .145833
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /map="20"
                     /clone="RP11-712N14"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .1214
                     /note="assembly_fragment"
     misc_feature    1315. .2673
                     /note="assembly_fragment"
     misc_feature    2774. .4520
                     /note="assembly_fragment"
     misc_feature    4621. .5961
                     /note="assembly_fragment"
     misc_feature    6062. .7719
                     /note="assembly_fragment"
     misc_feature    7820. .9799
                     /note="assembly_fragment"
     misc_feature    9900. .11434
                     /note="assembly_fragment"
     misc_feature    11535. .14382
                     /note="assembly_fragment"
     misc_feature    14483. .17148
                     /note="assembly_fragment"
     misc_feature    17249. .19239
                     /note="assembly_fragment"
     misc_feature    19340. .21102
                     /note="assembly_fragment"
     misc_feature    21203. .23371
                     /note="assembly_fragment"
     misc_feature    23472. .25782
                     /note="assembly_fragment"
     misc_feature    25883. .28824
                     /note="assembly_fragment"
     misc_feature    28925. .31619
                     /note="assembly_fragment"
     misc_feature    31720. .34679
                     /note="assembly_fragment"
     misc_feature    34780. .37527
                     /note="assembly_fragment"
     misc_feature    37628. .40146
                     /note="assembly_fragment"
     misc_feature    40247. .43743
                     /note="assembly_fragment"
     misc_feature    43844. .47133
                     /note="assembly_fragment"
     misc_feature    47234. .51023
                     /note="assembly_fragment"
     misc_feature    51124. .54935
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
     misc_feature    55036. .59553

  Query Match             100.0%;  Score 26;  DB 2;  Length 145833;
  Best Local Similarity   100.0%;  Pred. No. 0.21;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCTTCTGGAAGTCTGCCAGTGTCCTT 26
              ||||||||||||||||||||||||||
Db     100020 CCTTCTGGAAGTCTGCCAGTGTCCTT 99995


RESULT 9
AC129114/c
LOCUS       AC129114               177004 bp    DNA     linear   HTG 11-OCT-2002
DEFINITION  Rattus norvegicus clone CH230-251F22, WORKING DRAFT SEQUENCE, 4
            unordered pieces.
ACCESSION   AC129114
VERSION     AC129114.2  GI:23814684
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 177004)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
```

Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 177004)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (27-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 177004)
AUTHORS Rat Genome Sequencing Consortium.
TITLE Direct Submission
JOURNAL Submitted (11-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT On Oct 11, 2002 this sequence version replaced gi:21998900.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.
---------- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
---------- Project Information
Center project name: GTMI
Center clone name: CH230-251F22
---------- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 156282 bases at least Q40
Consensus quality: 158286 bases at least Q30
Consensus quality: 159515 bases at least Q20
Estimated insert size: 156592; sum-of-contigs estimation
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation
----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 173251: contig of 173251 bp in length
* 173252 173351: gap of unknown length
* 173352 174391: contig of 1040 bp in length
* 174392 174491: gap of unknown length
* 174492 175689: contig of 1198 bp in length
* 175690 175789: gap of unknown length
* 175790 177004: contig of 1215 bp in length.

FEATURES Location/Qualifiers
source 1. .177004
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-251F22"
misc_feature 1. .1427
/note="wgs_end_extension
clone_end:T7"
misc_feature 6478. .7966
/note="wgs_end_extension
clone_end:T7"
misc_feature 9311. .10240
/note="clone_boundary
clone_end:T7
site:MboI
end_sequence:RXABT35TJ"
misc_feature 54970. .165759
/note="clone_boundary
clone_end:Sp6
site:MboI
end_sequence:RXABT35TV"
misc_feature 171562. .173251
/note="wgs_end_extension
clone_end:Sp6"
ORIGIN

Query Match 81.5%; Score 21.2; DB 2; Length 177004;
Best Local Similarity 88.5%; Pred. No. 35;
Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CCTTCTGGAAGTCTGCCAGTGTCCTT 26
Db 95163 CTTTCAGGAATTCTGCCAGTGTCCTT 95138

RESULT 10
AC020636/c
LOCUS AC020636 179497 bp DNA linear PRI 15-MAR-2003
DEFINITION Homo sapiens 3 BAC RP11-251C9 (Roswell Park Cancer Institute Human BAC Library) complete sequence.
ACCESSION AC020636
VERSION AC020636.14 GI:22003934
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 179497)
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T., Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,

Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., He,X., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D., Nelson,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Naylor,S.L., Weinstock,G. and Gibbs,R.

TITLE Direct Submission
JOURNAL Unpublished

REFERENCE 2 (bases 1 to 179497)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (07-JAN-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 3 (bases 1 to 179497)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (26-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 4 (bases 1 to 179497)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (30-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 5 (bases 1 to 179497)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (24-JAN-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 6 (bases 1 to 179497)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (14-MAR-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 7 (bases 1 to 179497)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (15-MAR-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT On Jul 30, 2002 this sequence version replaced gi:21954832.
INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:
STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE:Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:
http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html.

FEATURES	Location/Qualifiers
source	1. .179497
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="3"
	/clone="RP11-251C9"
misc_feature	complement(1. .2007)
	/note="overlaps bases 1. .2007 of clone AC108674"
	/function="clone overlap"
repeat_region	complement(1719. .2026)
	/rpt_family="AluSx"
repeat_region	complement(2316. .2396)
	/rpt_family="MER91A"
STS	4464. .4739
	/standard_name="SHGC-80096"
repeat_region	6319. .6734
	/rpt_family="MLT2FB"
repeat_region	7371. .7400
	/rpt_family="AT_rich"
repeat_region	7744. .7892
	/rpt_family="MIR"
STS	7988. .8274
	/standard_name="SHGC-144786"
repeat_region	9667. .10430
	/rpt_family="MER21B"
repeat_region	complement(10508. .11039)
	/rpt_family="MLT1E1"
repeat_region	11100. .11266
	/rpt_family="MIR"
repeat_region	complement(11490. .11847)
	/rpt_family="MLT1K"
repeat_region	12962. .13005
	/rpt_family="MIR"
repeat_region	13873. .13953
	/rpt_family="MIR"
repeat_region	13992. .14099
	/rpt_family="MER58B"
repeat_region	14156. .14355
	/rpt_family="MER58B"
repeat_region	14760. .15104
	/rpt_family="L1MC5"
repeat_region	15370. .15397

```
                     /rpt_family="AT_rich"
     repeat_region   complement(15437. .15729)
                     /rpt_family="AluSc"
     repeat_region   complement(15863. .16265)
                     /rpt_family="MLT1J"
     repeat_region   16558. .16711
                     /rpt_family="MIR"
     STS             16714. .16885
                     /standard_name="G54351"
     repeat_region   complement(17909. .17988)
                     /rpt_family="MADE1"
     repeat_region   complement(20197. .20270)
                     /rpt_family="L2"
     repeat_region   21547. .21594
                     /rpt_family="(CA)n"
     repeat_region   21947. .22063
                     /rpt_family="(CA)n"
     repeat_region   complement(23769. .24118)
                     /rpt_family="MLT1J1"
     repeat_region   complement(27498. .27630)
                     /rpt_family="FLAM_C"
     repeat_region   27771. .27798
                     /rpt_family="AT_rich"
     repeat_region   27838. .27866
                     /rpt_family="(CA)n"
     repeat_region   29110. .29356
                     /rpt_family="MLT1J"
     repeat_region   29684. .29788
                     /rpt_family="MIR"

  Query Match             81.5%;  Score 21.2;  DB 9;  Length 179497;
  Best Local Similarity   88.5%;  Pred. No. 35;
  Matches   23;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CCTTCTGGAAGTCTGCCAGTGTCCTT 26
              || |||||||||| |||||||||  |
Db      29519 CCATCTGGAAGTCTCCCAGTGTCCAT 29494


RESULT 11
AC105625/c
LOCUS       AC105625               259967 bp    DNA     linear   HTG 19-NOV-2002
DEFINITION  Rattus norvegicus clone CH230-117017, WORKING DRAFT SEQUENCE, 2
            unordered pieces.
ACCESSION   AC105625
VERSION     AC105625.4  GI:25075238
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 259967)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 259967)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JAN-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 259967)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-NOV-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Nov 19, 2002 this sequence version replaced gi:23101542.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GNJS
                Center clone name: CH230-117017
            -------------- Summary Statistics
                Assembly program: Phrap; version 0.990329
                Consensus quality: 218447 bases at least Q40
                Consensus quality: 221126 bases at least Q30
                Consensus quality: 222950 bases at least Q20
                Estimated insert size: 224159; sum-of-contigs estimation
                Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
```

```
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 2 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    14629: contig of 14629 bp in length
            *    14630    14729: gap of unknown length
            *    14730   259967: contig of 245238 bp in length.
FEATURES             Location/Qualifiers
     source          1. .259967
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-117O17"
     misc_feature    1. .1404
                     /note="wgs_contig"
ORIGIN

  Query Match             81.5%;  Score 21.2;  DB 2;  Length 259967;
  Best Local Similarity   88.5%;  Pred. No. 34;
  Matches   23;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CCTTCTGGAAGTCTGCCAGTGTCCTT 26
              | ||| |||| ||||||||||||||||
Db     168681 CTTTCAGGAATTCTGCCAGTGTCCTT 168656


RESULT 12
AC095338/c
LOCUS       AC095338               261492 bp    DNA     linear   HTG 10-MAY-2003
DEFINITION  Rattus norvegicus clone CH230-114A11, *** SEQUENCING IN PROGRESS
            ***, 7 unordered pieces.
ACCESSION   AC095338
VERSION     AC095338.6  GI:30522611
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 261492)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 261492)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-SEP-2001) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 261492)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On May 10, 2003 this sequence version replaced gi:23096505.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            ---------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            ---------- Project Information
                Center project name: GGNW
                Center clone name: CH230-114A11
            ---------- Summary Statistics
                Assembly program: Atlas 3.0;
                Consensus quality: 244697 bases at least Q40
                Consensus quality: 248589 bases at least Q30
                Consensus quality: 251288 bases at least Q20
                Estimated insert size: 255184; sum-of-contigs estimation
                Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
            ----------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
            * NOTE: This sequence may represent more than one clone.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 7 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
```

```
            * as soon as it is available and the accession number will
            * be preserved.
            *         1   249471: contig of 249471 bp in length
            *    249472   249571: gap of unknown length
            *    249572   250636: contig of 1065 bp in length
            *    250637   250736: gap of unknown length
            *    250737   251874: contig of 1138 bp in length
            *    251875   251974: gap of unknown length
            *    251975   253893: contig of 1919 bp in length
            *    253894   253993: gap of unknown length
            *    253994   255314: contig of 1321 bp in length
            *    255315   255414: gap of unknown length
            *    255415   256732: contig of 1318 bp in length
            *    256733   256832: gap of unknown length
            *    256833   261492: contig of 4660 bp in length.
FEATURES             Location/Qualifiers
     source          1. .261492
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-114A11"
     misc_feature    1. .1458
                     /note="wgs_contig"
     misc_feature    complement(247382. .248063)
                     /note="clone_boundary
                     clone_end:T7
                     site:EcoRI
                     end_sequence:BH288027"
ORIGIN

  Query Match             81.5%;  Score 21.2;  DB 2;  Length 261492;
  Best Local Similarity   88.5%;  Pred. No. 34;
  Matches   23;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CCTTCTGGAAGTCTGCCAGTGTCCTT 26
              |||| |||| ||||||||||||||||
Db      52710 CTTTCAGGAATTCTGCCAGTGTCCTT 52685


RESULT 13
AC021312
LOCUS       AC021312               76332 bp    DNA     linear   HTG 13-JUL-2000
DEFINITION  Homo sapiens chromosome 4 clone RP11-197L13 map 4, LOW-PASS
            SEQUENCE SAMPLING.
ACCESSION   AC021312
VERSION     AC021312.2  GI:9126405
KEYWORDS    HTG; HTGS_PHASE0.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 76332)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 4, clone RP11-197L13
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 76332)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
            DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
            Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
            Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Landers,T., Lehoczky,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
            Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
            McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
            Norman,C.H., O'Connor,T., O'Donnell,P., Olivar,T.M., Peterson,K.,
            Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
            Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
            Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Jul 13, 2000 this sequence version replaced gi:6705794.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L4256
                Center clone name: 197_L_13
            --------------.
            * NOTE: This record contains 84 individual
            * sequencing reads that have not been assembled into
            * contigs. Runs of N are used to separate the reads
            * and the order in which they appear is completely
            * arbitrary. Low-pass sequence sampling is useful for
            * identifying clones that may be gene-rich and allows
            * overlap relationships among clones to be deduced.
            * However, it should not be assumed that this clone
            * will be sequenced to completion. In the event that
            * the record is updated, the accession number will
            * be preserved.
            *         1     809: contig of 809 bp in length
            *       810     909: gap of 100 bp
            *       910    1713: contig of 804 bp in length
            *      1714    1813: gap of 100 bp
            *      1814    2627: contig of 814 bp in length
            *      2628    2727: gap of 100 bp
            *      2728    3523: contig of 796 bp in length
            *      3524    3623: gap of 100 bp
            *      3624    4432: contig of 809 bp in length
            *      4433    4532: gap of 100 bp
            *      4533    5324: contig of 792 bp in length
            *      5325    5424: gap of 100 bp
            *      5425    6207: contig of 783 bp in length
            *      6208    6307: gap of 100 bp
            *      6308    7094: contig of 787 bp in length
            *      7095    7194: gap of 100 bp
            *      7195    8019: contig of 825 bp in length
            *      8020    8119: gap of 100 bp
            *      8120    8909: contig of 790 bp in length
            *      8910    9009: gap of 100 bp
            *      9010    9839: contig of 830 bp in length
            *      9840    9939: gap of 100 bp
            *      9940   10753: contig of 814 bp in length
            *     10754   10853: gap of 100 bp
            *     10854   11665: contig of 812 bp in length
            *     11666   11765: gap of 100 bp
            *     11766   12577: contig of 812 bp in length
            *     12578   12677: gap of 100 bp
            *     12678   13494: contig of 817 bp in length
            *     13495   13594: gap of 100 bp
            *     13595   14384: contig of 790 bp in length
            *     14385   14484: gap of 100 bp
            *     14485   15300: contig of 816 bp in length
            *     15301   15400: gap of 100 bp
            *     15401   16201: contig of 801 bp in length
            *     16202   16301: gap of 100 bp
            *     16302   17092: contig of 791 bp in length
            *     17093   17192: gap of 100 bp
            *     17193   17983: contig of 791 bp in length
            *     17984   18083: gap of 100 bp
            *     18084   18890: contig of 807 bp in length
            *     18891   18990: gap of 100 bp
            *     18991   19817: contig of 827 bp in length
            *     19818   19917: gap of 100 bp
            *     19918   20737: contig of 820 bp in length
            *     20738   20837: gap of 100 bp
```

```
*    20838    21660: contig of 823 bp in length
*    21661    21760: gap of 100 bp
*    21761    22575: contig of 815 bp in length
*    22576    22675: gap of 100 bp
*    22676    23466: contig of 791 bp in length
*    23467    23566: gap of 100 bp
*    23567    24371: contig of 805 bp in length
*    24372    24471: gap of 100 bp
*    24472    25253: contig of 782 bp in length
*    25254    25353: gap of 100 bp
*    25354    26176: contig of 823 bp in length
*    26177    26276: gap of 100 bp
*    26277    27096: contig of 820 bp in length
*    27097    27196: gap of 100 bp
*    27197    28022: contig of 826 bp in length
*    28023    28122: gap of 100 bp
*    28123    28937: contig of 815 bp in length
*    28938    29037: gap of 100 bp
*    29038    29856: contig of 819 bp in length
*    29857    29956: gap of 100 bp
*    29957    30764: contig of 808 bp in length
*    30765    30864: gap of 100 bp
*    30865    31672: contig of 808 bp in length
*    31673    31772: gap of 100 bp
*    31773    32594: contig of 822 bp in length
*    32595    32694: gap of 100 bp
*    32695    33471: contig of 777 bp in length
*    33472    33571: gap of 100 bp
*    33572    34419: contig of 848 bp in length
*    34420    34519: gap of 100 bp
*    34520    35349: contig of 830 bp in length
*    35350    35449: gap of 100 bp
*    35450    36266: contig of 817 bp in length
*    36267    36366: gap of 100 bp
*    36367    37180: contig of 814 bp in length
*    37181    37280: gap of 100 bp
*    37281    38094: contig of 814 bp in length
*    38095    38194: gap of 100 bp
*    38195    39006: contig of 812 bp in length
*    39007    39106: gap of 100 bp
*    39107    39915: contig of 809 bp in length
*    39916    40015: gap of 100 bp
*    40016    40825: contig of 810 bp in length
*    40826    40925: gap of 100 bp
*    40926    41733: contig of 808 bp in length
*    41734    41833: gap of 100 bp
*    41834    42620: contig of 787 bp in length
*    42621    42720: gap of 100 bp
*    42721    43510: contig of 790 bp in length
*    43511    43610: gap of 100 bp
*    43611    44396: contig of 786 bp in length
*    44397    44496: gap of 100 bp
*    44497    45321: contig of 825 bp in length
*    45322    45421: gap of 100 bp
*    45422    46237: contig of 816 bp in length
*    46238    46337: gap of 100 bp
*    46338    47160: contig of 823 bp in length
*    47161    47260: gap of 100 bp
*    47261    48063: contig of 803 bp in length
*    48064    48163: gap of 100 bp
*    48164    48971: contig of 808 bp in length
*    48972    49071: gap of 100 bp
*    49072    49901: contig of 830 bp in length
*    49902    50001: gap of 100 bp
*    50002    50794: contig of 793 bp in length
*    50795    50894: gap of 100 bp
*    50895    51725: contig of 831 bp in length
*    51726    51825: gap of 100 bp
*    51826    52619: contig of 794 bp in length
*    52620    52719: gap of 100 bp
*    52720    53496: contig of 777 bp in length
*    53497    53596: gap of 100 bp
*    53597    54412: contig of 816 bp in length
*    54413    54512: gap of 100 bp
*    54513    55337: contig of 825 bp in length
*    55338    55437: gap of 100 bp
*    55438    56254: contig of 817 bp in length
*    56255    56354: gap of 100 bp
*    56355    57160: contig of 806 bp in length
*    57161    57260: gap of 100 bp
*    57261    58081: contig of 821 bp in length
*    58082    58181: gap of 100 bp
*    58182    58996: contig of 815 bp in length
*    58997    59096: gap of 100 bp
*    59097    59896: contig of 800 bp in length
*    59897    59996: gap of 100 bp
*    59997    60802: contig of 806 bp in length
*    60803    60902: gap of 100 bp
*    60903    61690: contig of 788 bp in length
*    61691    61790: gap of 100 bp
*    61791    62590: contig of 800 bp in length
*    62591    62690: gap of 100 bp
*    62691    63480: contig of 790 bp in length
*    63481    63580: gap of 100 bp
*    63581    64395: contig of 815 bp in length

  Query Match             80.0%;  Score 20.8;  DB 2;  Length 76332;
  Best Local Similarity   91.7%;  Pred. No. 59;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 CTTCTGGAAGTCTGCCAGTGTCCT 25
              |||||||||||||||||||| | |
Db      70153 CTTCTGGAAGTCTGCCAGTGGCTT 70176


RESULT 14
AC011405/c
LOCUS       AC011405               184541 bp    DNA     linear   PRI 27-FEB-2002
DEFINITION  Homo sapiens chromosome 5 clone CTB-46B19, complete sequence.
ACCESSION   AC011405
VERSION     AC011405.6  GI:18921282
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 184541)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 184541)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-OCT-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 184541)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-FEB-2002) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
REFERENCE   4  (bases 1 to 184541)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-FEB-2002) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Feb 26, 2002 this sequence version replaced gi:8576061.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 100% of Sequence;
            Estimated Total Number of Errors is 0.
            NOTE:  BACTERIAL TRANSPOSON excised at 162460
            NOTE:  This insert is not the entire sequence of the clone (entire
            sequence is 195.7kb). It is clipped at the overlap with AC034243.
            The number of bases overlapped is 32318.
```

```
FEATURES             Location/Qualifiers
     source          1. .184541
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="CTB-46B19"
ORIGIN

  Query Match             80.0%;  Score 20.8;  DB 9;  Length 184541;
  Best Local Similarity   91.7%;  Pred. No. 54;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 CTTCTGGAAGTCTGCCAGTGTCCT 25
              |||||||||||||||||||| | |
Db     108454 CTTCTGGAAGTCTGCCAGTGGCTT 108431


RESULT 15
AL355474/c
LOCUS       AL355474              120194 bp    DNA     linear   PRI 09-JAN-2002
DEFINITION  Human DNA sequence from clone RP11-55C5 on chromosome 10, complete
            sequence.
ACCESSION   AL355474
VERSION     AL355474.15  GI:18121463
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 120194)
  AUTHORS   Almeida,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JAN-2002) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Jan 10, 2002 this sequence version replaced gi:18072474.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em:, EMBL; Sw:,
            SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 10, constructed by the Sanger Centre Chromosome 10
            Mapping Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr10
            RP11-55C5 is from the library RPCI-11.1 constructed by the group of
            Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            IMPORTANT: This sequence is not the entire insert of clone
            RP11-55C5 It may be shorter because we sequence overlapping
            sections only once, except for a short overlap.
            The true left end of clone RP11-55C5 is at 1 in this sequence. The
            true left end of clone RP11-357A18 is at 118195 in this sequence.
FEATURES             Location/Qualifiers
     source          1. .120194
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="10"
                     /clone="RP11-55C5"
                     /clone_lib="RPCI-11.1"
     misc_feature    39283. .39323
                     /note="Sequence from AC025038 sequenced by GTC."
ORIGIN

  Query Match             78.5%;  Score 20.4;  DB 9;  Length 120194;
  Best Local Similarity   95.5%;  Pred. No. 87;
  Matches   21;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 CTTCTGGAAGTCTGCCAGTGTC 23
              |||||||||||||||||||| |
Db      71803 CTTCTGGAAGTCTGCCAGTGCC 71782


Search completed: February 19, 2004, 23:23:23
Job time : 119.831 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 21:15:48 ; Search time 24.0879 Seconds
(without alignments)
4585.415 Million cell updates/sec

Title: US-09-939-853A-141
Perfect score: 26
Sequence: 1 ccttctggaagtctgccagtgtcctt 26

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	26	100.0	26	6	ABK61506	Abk61506 Human NOV
c	2	26	100.0	444	6	ABQ98669	Abq98669 Human ORF
	3	26	100.0	445	5	AAS74747	Aas74747 DNA encod
c	4	26	100.0	763	6	ABQ98670	Abq98670 Human ORF
c	5	26	100.0	1183	6	ABK61465	Abk61465 Human cDN
c	6	26	100.0	2567	6	AAD43980	Aad43980 Human Src
	7	18.8	72.3	258	8	ACC58334	Acc58334 Human abr
	8	18.8	72.3	258	8	ACC58338	Acc58338 Human abr
	9	18.8	72.3	258	8	ACC58337	Acc58337 Mouse abr
	10	18.8	72.3	258	8	ACC58336	Acc58336 Human abr
	11	18.8	72.3	264	6	AAD27075	Aad27075 Human uro
	12	18.8	72.3	288	6	AAD27083	Aad27083 Human uPA
	13	18.8	72.3	300	1	AAN81299	Aan81299 Human pro
	14	18.8	72.3	405	6	AAD27078	Aad27078 Human uro
	15	18.8	72.3	429	6	AAD27082	Aad27082 Human uPA
	16	18.8	72.3	465	7	AAL55230	Aal55230 Tumour an
	17	18.8	72.3	465	7	AAL55236	Aal55236 Tumour an
	18	18.8	72.3	482	6	ABK62674	Abk62674 Rat seque
	19	18.8	72.3	482	9	ADB56463	Adb56463 Toxicity-
	20	18.8	72.3	482	9	ADB50966	Adb50966 Primary r
	21	18.8	72.3	618	9	ADD35169	Add35169 Mouse mit
	22	18.8	72.3	624	2	AAT75153	Aat75153 Metastasi
	23	18.8	72.3	645	2	AAT75154	Aat75154 Metastasi
	24	18.8	72.3	666	2	AAT75155	Aat75155 Metastasi
c	25	18.8	72.3	689	6	ABI99589	Abi99589 Mouse isc
	26	18.8	72.3	1137	2	AAT61672	Aat61672 cDNA enco
	27	18.8	72.3	1170	2	AAT61673	Aat61673 Human pro
	28	18.8	72.3	1188	2	AAQ55771	Aaq55771 Pro-uroki
	29	18.8	72.3	1205	2	AAT61674	Aat61674 Human pro
	30	18.8	72.3	1212	6	AAD27080	Aad27080 Human uro
	31	18.8	72.3	1233	2	AAT18237	Aat18237 Pro-uroki
	32	18.8	72.3	1233	2	AAT68797	Aat68797 Coding se
	33	18.8	72.3	1236	2	AAQ06133	Aaq06133 Sequence
	34	18.8	72.3	1236	2	AAQ06134	Aaq06134 Sequence
	35	18.8	72.3	1236	2	AAQ06135	Aaq06135 Sequence
	36	18.8	72.3	1236	2	AAQ04486	Aaq04486 Plasmid p
	37	18.8	72.3	1236	2	AAQ04485	Aaq04485 Plasmid p
	38	18.8	72.3	1236	2	AAT61671	Aat61671 Human nat
	39	18.8	72.3	1236	2	AAQ10168	Aaq10168 Encodes P
	40	18.8	72.3	1236	2	AAQ10169	Aaq10169 Encodes P
	41	18.8	72.3	1236	2	AAQ10170	Aaq10170 Encodes P
	42	18.8	72.3	1236	2	AAQ23009	Aaq23009 Pro-UK. 3
	43	18.8	72.3	1236	2	AAQ41450	Aaq41450 Mutant hu
	44	18.8	72.3	1236	2	AAQ48228	Aaq48228 PUK gene.
	45	18.8	72.3	1236	6	AAD27077	Aad27077 Human uro

ALIGNMENTS

RESULT 1
ABK61506
ID ABK61506 standard; DNA; 26 BP.
XX
AC ABK61506;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human NOV13 RT-PCR probe.
XX
KW Human; ss; NOVX; gene therapy; cardiomyopathy; atherosclerosis;
KW cell signal processing disorder; metabolic pathway modulation disorder;
KW diabetes; cancer; adenocarcinoma; lymphoma; prostate cancer; probe;
KW uterus cancer; immune response; graft-versus-host disease; Exon linking;
KW acquired immunodeficiency syndrome; AIDS; asthma; Crohn's disease;
KW hypertension; congenital heart defects; multiple sclerosis; inflammation;
KW Albright hereditary osteodystrophy; reverse transcriptase PCR.
XX
OS Homo sapiens.
XX
PN WO200216599-A2.
XX
PD 28-FEB-2002.
XX
PF 27-AUG-2001; 2001WO-US026510.
XX
PR 25-AUG-2000; 2000US-0228191P.
PR 08-FEB-2001; 2001US-0267300P.
PR 20-FEB-2001; 2001US-0269961P.
PR 20-MAR-2001; 2001US-0277337P.
XX
PA (CURA-) CURAGEN CORP.
PA (CORT-) COR THERAPEUTICS INC.
XX
PI Burgess CE, Conley PB, Grosse WM, Hart M, Kekuda R, Shimkets RA;
PI Spytek KA, Szekeres ES, Tomlinson JE, Topper JN, Yang R;
XX
DR WPI; 2002-280937/32.
XX
PT New polypeptides for treating or preventing a disorder associated with
PT them, in humans, e.g. cardiomyopathy, atherosclerosis or cancers.
XX
PS Example 2; Page 234; 263pp; English.
XX
CC The invention relates to an isolated polypeptide (NOVX) a mature form of
CC NOVX, a NOVX variant (differing by no more than 15%), the nucleotide

CC encoding NOVX (or its complement, fragment or variant). NOVX is NOV1-14,
CC 15a, 15b, 16a, and 16b. The NOVX polypeptide, nucleic acid encoding it
CC and antibody against it, are useful for treating or preventing (e.g. by
CC gene therapy) a NOVX-associated disorder in humans, e.g. cardiomyopathy,
CC atherosclerosis, a disorder related to cell signal processing and
CC metabolic pathway modulation, diabetes or cancers. The NOVX polypeptide
CC and nucleic acids are also useful for determining the presence of
CC predisposition to the diseases. The NOVX nucleic acid and polypeptide are
CC especially useful in therapeutic or prophylactic applications for
CC disorders associated with aberrant NOVX expression or activity, e.g.
CC cancers (e.g. adenocarcinoma, lymphoma, prostate cancer or uterus
CC cancer), immune response, graft-versus-host disease, acquired
CC immunodeficiency syndrome (AIDS), asthma, Crohn's disease, hypertension,
CC congenital heart defects, multiple sclerosis, inflammation or Albright
CC hereditary osteodystrophy and many other diseases listed in the
CC specification. The DNA encoding the protein is useful in gene therapy for
CC treating the conditions. This is also useful in detection assays,
CC chromosome mapping, tissue typing, diagnostic or prognostic assays, or
CC for developing a powerful assay system for functional analysis of various
CC human disorders, as well as in diagnostic applications. The present
CC sequence is a reverse transcriptase (RT)-PCR probe used to measure tissue
CC specific expression of mRNA encoding a NOVX protein
XX
SQ Sequence 26 BP; 3 A; 8 C; 6 G; 9 T; 0 U; 0 Other;

Query Match 100.0%; Score 26; DB 6; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.064;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCTTCTGGAAGTCTGCCAGTGTCCTT 26
||||||||||||||||||||||||||
Db 1 CCTTCTGGAAGTCTGCCAGTGTCCTT 26

RESULT 2
ABQ98669/c
ID ABQ98669 standard; DNA; 444 BP.
XX
AC ABQ98669;
XX
DT 04-NOV-2002 (first entry)
XX
DE Human ORF476 coding sequence.
XX
KW Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnerary;
KW Antiinflammatory; gene therapy; human; ORFX; atherogenic; platelet;
KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;
KW cancer; cardiovascular disease; allergy; autoimmune disease;
KW wound healing; blood coagulation disorder; inflammatory disorder; ds.
XX
OS Homo sapiens.
XX
PN US2002082206-A1.
XX
PD 27-JUN-2002.
XX
PF 30-MAY-2001; 2001US-00867550.
XX
PR 30-MAY-2000; 2000US-0208427P.
XX
PA (LEAC/) LEACH M D.
PA (MEHR/) MEHRABAN F.
PA (CONL/) CONLEY P B.
PA (TOPP/) TOPPER J N.
PA (LAWD/) LAW D.
XX
PI Leach MD, Mehraban F, Conley PB, Topper JN, Law D;
XX
DR WPI; 2002-626554/67.
DR P-PSDB; ABP64106.
XX
PT New polypeptide designated ORFX are present in human atherogenic cells
PT and are useful to prevent and treat ORFX-associated disorders including
PT cancer, allergy, wound healing or autoimmune, cardiovascular or
PT inflammatory disease.
XX
PS Claim 2; SEQ ID NO 951; 78pp; English.
XX
CC The present invention relates to novel human ORFX polypeptides and their
CC coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences
CC were discovered in human atherogenic cells, in particular in platelets
CC and human umbilical vein endothelial cells (HUVEC) and are expressed in
CC many other tissues as well. Atherogenic cells are cells which have the
CC potential to develop atherosclerotic plaques. The ORFX polypeptides and
CC nucleic acids are useful for treating or preventing a pathological
CC condition associated with an ORFX-associated disorder, e.g. cancer,
CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood
CC coagulation disorders or inflammatory disorders. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from the USPTO web site at
CC seqdata.uspto.gov/sequence.html?DocID=20020082206
XX
SQ Sequence 444 BP; 103 A; 128 C; 132 G; 81 T; 0 U; 0 Other;

Query Match 100.0%; Score 26; DB 6; Length 444;
Best Local Similarity 100.0%; Pred. No. 0.09;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCTTCTGGAAGTCTGCCAGTGTCCTT 26
||||||||||||||||||||||||||
Db 35 CCTTCTGGAAGTCTGCCAGTGTCCTT 10

RESULT 3
AAS74747
ID AAS74747 standard; cDNA; 445 BP.
XX
AC AAS74747;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #10551.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG10560.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 10551; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed

CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 445 BP; 89 A; 112 C; 143 G; 101 T; 0 U; 0 Other;

Query Match 100.0%; Score 26; DB 5; Length 445;
Best Local Similarity 100.0%; Pred. No. 0.09;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCTTCTGGAAGTCTGCCAGTGTCCTT 26
||||||||||||||||||||||||||
Db 193 CCTTCTGGAAGTCTGCCAGTGTCCTT 218

RESULT 4
ABQ98670/c
ID ABQ98670 standard; DNA; 763 BP.
XX
AC ABQ98670;
XX
DT 04-NOV-2002 (first entry)
XX
DE Human ORF477 coding sequence.
XX
KW Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnerary;
KW Antiinflammatory; gene therapy; human; ORFX; atherogenic; platelet;
KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;
KW cancer; cardiovascular disease; allergy; autoimmune disease;
KW wound healing; blood coagulation disorder; inflammatory disorder; ds.
XX
OS Homo sapiens.
XX
PN US2002082206-A1.
XX
PD 27-JUN-2002.
XX
PF 30-MAY-2001; 2001US-00867550.
XX
PR 30-MAY-2000; 2000US-0208427P.
XX
PA (LEAC/) LEACH M D.
PA (MEHR/) MEHRABAN F.
PA (CONL/) CONLEY P B.
PA (TOPP/) TOPPER J N.
PA (LAWD/) LAW D.
XX
PI Leach MD, Mehraban F, Conley PB, Topper JN, Law D;
XX
DR WPI; 2002-626554/67.
DR P-PSDB; ABP64107.
XX
PT New polypeptide designated ORFX are present in human atherogenic cells
PT and are useful to prevent and treat ORFX-associated disorders including
PT cancer, allergy, wound healing or autoimmune, cardiovascular or
PT inflammatory disease.
XX
PS Claim 2; SEQ ID NO 953; 78pp; English.
XX
CC The present invention relates to novel human ORFX polypeptides and their
CC coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences
CC were discovered in human atherogenic cells, in particular in platelets
CC and human umbilical vein endothelial cells (HUVEC) and are expressed in
CC many other tissues as well. Atherogenic cells are cells which have the
CC potential to develop atherosclerotic plaques. The ORFX polypeptides and
CC nucleic acids are useful for treating or preventing a pathological
CC condition associated with an ORFX-associated disorder, e.g. cancer,
CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood
CC coagulation disorders or inflammatory disorders. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from the USPTO web site at
CC seqdata.uspto.gov/sequence.html?DocID=20020082206
XX
SQ Sequence 763 BP; 176 A; 222 C; 218 G; 147 T; 0 U; 0 Other;

Query Match 100.0%; Score 26; DB 6; Length 763;
Best Local Similarity 100.0%; Pred. No. 0.096;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCTTCTGGAAGTCTGCCAGTGTCCTT 26
||||||||||||||||||||||||||
Db 164 CCTTCTGGAAGTCTGCCAGTGTCCTT 139

RESULT 5
ABK61465/c
ID ABK61465 standard; cDNA; 1183 BP.
XX
AC ABK61465;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human cDNA encoding protein NOV13.
XX
KW Human; gene; ss; NOVX; gene therapy; cardiomyopathy; atherosclerosis;
KW cell signal processing disorder; metabolic pathway modulation disorder;
KW diabetes; cancer; adenocarcinoma; lymphoma; prostate cancer;
KW uterus cancer; immune response; graft-versus-host disease;
KW acquired immunodeficiency syndrome; AIDS; asthma; Crohn's disease;
KW hypertension; congenital heart defects; multiple sclerosis; inflammation;
KW Albright hereditary osteodystrophy.
XX
OS Homo sapiens.
XX
PN WO200216599-A2.
XX
PD 28-FEB-2002.
XX
PF 27-AUG-2001; 2001WO-US026510.
XX
PR 25-AUG-2000; 2000US-0228191P.
PR 08-FEB-2001; 2001US-0267300P.
PR 20-FEB-2001; 2001US-0269961P.
PR 20-MAR-2001; 2001US-0277337P.
XX
PA (CURA-) CURAGEN CORP.
PA (CORT-) COR THERAPEUTICS INC.
XX
PI Burgess CE, Conley PB, Grosse WM, Hart M, Kekuda R, Shimkets RA;
PI Spytek KA, Szekeres ES, Tomlinson JE, Topper JN, Yang R;
XX
DR WPI; 2002-280937/32.
DR P-PSDB; AAU91308.
XX
PT New polypeptides for treating or preventing a disorder associated with
PT them, in humans, e.g. cardiomyopathy, atherosclerosis or cancers.
XX
PS Claim 1; Page 98; 263pp; English.
XX
CC The invention relates to an isolated polypeptide (NOVX) a mature form of
CC NOVX, a NOVX variant (differing by no more than 15%), the nucleotide
CC encoding NOVX (or its complement, fragment or variant). NOVX is NOV1-14,
CC 15a, 15b, 16a, and 16b. The NOVX polypeptide, nucleic acid encoding it

```
CC   and antibody against it, are useful for treating or preventing (e.g. by
CC   gene therapy) a NOVX-associated disorder in humans, e.g. cardiomyopathy,
CC   atherosclerosis, a disorder related to cell signal processing and
CC   metabolic pathway modulation, diabetes or cancers. The NOVX polypeptide
CC   and nucleic acids are also useful for determining the presence of
CC   predisposition to the diseases. The NOVX nucleic acid and polypeptide are
CC   especially useful in therapeutic or prophylactic applications for
CC   disorders associated with aberrant NOVX expression or activity, e.g.
CC   cancers (e.g. adenocarcinoma, lymphoma, prostate cancer or uterus
CC   cancer), immune response, graft-versus-host disease, acquired
CC   immunodeficiency syndrome (AIDS), asthma, Crohn's disease, hypertension,
CC   congenital heart defects, multiple sclerosis, inflammation or Albright
CC   hereditary osteodystrophy and many other diseases listed in the
CC   specification. The DNA encoding the protein is useful in gene therapy for
CC   treating the conditions. This is also useful in detection assays,
CC   chromosome mapping, tissue typing, diagnostic or prognostic assays, or
CC   for developing a powerful assay system for functional analysis of various
CC   human disorders, as well as in diagnostic applications. The present
CC   sequence encodes a NOVX protein
XX
SQ   Sequence 1183 BP; 251 A; 359 C; 333 G; 240 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 26;  DB 6;  Length 1183;
  Best Local Similarity   100.0%;  Pred. No. 0.1;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCTTCTGGAAGTCTGCCAGTGTCCTT 26
              ||||||||||||||||||||||||||
Db        276 CCTTCTGGAAGTCTGCCAGTGTCCTT 251


RESULT 6
AAD43980/c
ID   AAD43980 standard; cDNA; 2567 BP.
XX
AC   AAD43980;
XX
DT   13-DEC-2002  (first entry)
XX
DE   Human Src-Like Adapter Protein-2 (hSLAP-2) cDNA.
XX
KW   Human; SH2/SH3-domain-containing adapter; Src-Like Adapter Protein-2;
KW   SLAP-2; immune disorder; signal transduction; autoimmune disease; cancer;
KW   neoplasm; inflammation disorder; rheumatoid arthritis; osteoarthritis;
KW   psoriasis; rhinitis; allergy; inflammatory bowel disease; gene therapy;
KW   Crohn's disease; systemic lupus erythematosus; tissue/organ rejection;
KW   multiple sclerosis; asthma; acute respiratory distress syndrome;
KW   pulmonary disorder; dermatological; neuroprotective; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             415. .1200
FT                   /*tag=  a
FT                   /product= "Human SLAP-2"
XX
PN   WO200242457-A1.
XX
PD   30-MAY-2002.
XX
PF   20-NOV-2001; 2001WO-US043367.
XX
PR   22-NOV-2000; 2000US-0252545P.
XX
PA   (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
PI   Chang H,  Yang W,  Wu Y,  Whitney GS,  Perez-Villar JJ,  Kanner SB;
XX
DR   WPI; 2002-463632/49.
DR   P-PSDB; AAE26357.
XX
PT   Novel substantially purified human SH2/SH3-domain-containing adapter
PT   polypeptide, termed Src-Like Adapter Protein-2, useful for therapeutic
PT   intervention in immunological and inflammatory disorders and cancer.
XX
PS   Claim 2; Fig 1; 85pp; English.
XX
CC   The invention relates to a substantially purified human SH2/SH3-domain-
CC   containing adapter polypeptide, termed Src-Like Adapter Protein-2 (SLAP-
CC   2). The invention is useful for treating an immune disorder involving
CC   hyperactivity of B- or T- lymphocytes in a mammal. The invention is
CC   useful for screening for antagonists or inhibitors of the interaction of
CC   hSLAP-2 with cellular signalling compounds, for diagnosing, treating or
CC   preventing diseases or disorders associated with aberrant or uncontrolled
CC   cellular signal transduction, for determining those cellular signalling
CC   molecules which associate with hSLAP-2 and which provide critical signals
CC   for cell activation, and as effectors in methods to affect T- cell
CC   activation. The invention is useful in screening assays to identity and
CC   detect candidate bioactive agents that modulate hSLAP-2 bioactivity, for
CC   potential use to treat autoimmune diseases which may be caused by
CC   hyperactivated B cells, as well as to treat diseases which may be caused
CC   by hyperactivated T cells, in addition to other immune system related
CC   conditions, diseases, or disorders, T-cell and B-cell neoplasms,
CC   inflammation disorders, diseases and conditions, rheumatoid arthritis,
CC   osteoarthritis, psoriasis, rhinitis, inflammatory bowel disease (Crohn's
CC   and ulcerative colitis), allergies, particularly those involving
CC   hyperactivity of B-cells and T- cells, or other immune cells, such as
CC   mast cells or eosinophils, autoimmune diseases such as systemic lupus
CC   erythematosus and multiple sclerosis, pulmonary diseases including
CC   asthma, acute respiratory distress syndrome, and chronic obstructive
CC   pulmonary disorder, tissue/ organ rejection and cancer. The invention is
CC   useful in gene therapy. The present sequence is human SLAP-2 cDNA
XX
SQ   Sequence 2567 BP; 611 A; 741 C; 666 G; 549 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 26;  DB 6;  Length 2567;
  Best Local Similarity   100.0%;  Pred. No. 0.11;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCTTCTGGAAGTCTGCCAGTGTCCTT 26
              ||||||||||||||||||||||||||
Db        293 CCTTCTGGAAGTCTGCCAGTGTCCTT 268


RESULT 7
ACC58334
ID   ACC58334 standard; cDNA; 258 BP.
XX
AC   ACC58334;
XX
DT   26-AUG-2003  (first entry)
XX
DE   Human abrogen (hATF-kringle) coding sequence.
XX
KW   Human; abrogen; kringle; hATF-kringle; angiogenesis; inhibitor;
KW   urokinase plasminogen activator; tumour; metastasis; cytostatic;
KW   gene therapy; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .258
FT                   /*tag=  a
FT                   /partial
FT                   /product= "Human abrogen"
FT                   /note= "no start or stop codon"
XX
PN   WO2003042354-A2.
XX
PD   22-MAY-2003.
XX
PF   04-SEP-2002; 2002WO-US027885.
XX
PR   04-SEP-2001; 2001US-0316300P.
XX
PA   (AVET ) AVENTIS PHARM INC.
```

```
XX
PI   Nesbit M,  Fong TC,  Brockstedt D;
XX
DR   WPI; 2003-449566/42.
DR   P-PSDB; ABR42599.
XX
PT   New abrogen polypeptide, useful for treating an angiogenesis related
PT   diseases e.g. tumor metastasis.
XX
PS   Claim 32; Page 24-25; 95pp; English.
XX
CC   The present sequence is a coding sequence for a novel human abrogen,
CC   designated hATF-kringle, comprising the human urokinase plasminogen
CC   activator kringle domain. Abrogens such as hATF-kringle are potent
CC   inhibitors of endothelial proliferation and angiogenesis. Abrogen
CC   polypeptides are capable of inhibiting or reducing cell proliferation
CC   induced by both basic fibroblast growth factor (bFGF) and vascular
CC   endothelial growth factor in a specific endothelial cell proliferation
CC   assay; angiostatin only inhibits bFGF induced proliferation in this
CC   assay. Vectors that expressed abrogen polypeptides in vivo were shown to
CC   reduce tumour metastasis in 2 lung cancer models. The invention provides
CC   abrogen polypeptides and polynucleotides, and methods of using these to
CC   treat an angiogenesis-related disease or disorder, e.g. tumour metastasis
CC   (claimed)
XX
SQ   Sequence 258 BP; 63 A; 74 C; 71 G; 50 T; 0 U; 0 Other;

  Query Match             72.3%;  Score 18.8;  DB 8;  Length 258;
  Best Local Similarity   90.9%;  Pred. No. 1.3e+02;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 CTGGAAGTCTGCCAGTGTCCTT 26
              |||||| ||||||| |||||||
Db         78 CTGGAACTCTGCCACTGTCCTT 99


RESULT 8
ACC58338
ID   ACC58338 standard; cDNA; 258 BP.
XX
AC   ACC58338;
XX
DT   26-AUG-2003  (first entry)
XX
DE   Human abrogen (hATF-kringle) coding sequence.
XX
KW   Human; abrogen; kringle; hATF-kringle; angiogenesis; inhibitor;
KW   urokinase plasminogen activator; tumour; metastasis; cytostatic;
KW   gene therapy; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .258
FT                   /*tag=  a
FT                   /partial
FT                   /product= "Human abrogen"
FT                   /note= "no start or stop codon"
XX
PN   WO2003042354-A2.
XX
PD   22-MAY-2003.
XX
PF   04-SEP-2002; 2002WO-US027885.
XX
PR   04-SEP-2001; 2001US-0316300P.
XX
PA   (AVET ) AVENTIS PHARM INC.
XX
PI   Nesbit M,  Fong TC,  Brockstedt D;
XX
DR   WPI; 2003-449566/42.
DR   P-PSDB; ABR42617.
XX
PT   New abrogen polypeptide, useful for treating an angiogenesis related
PT   diseases e.g. tumor metastasis.
XX
PS   Disclosure; Page 95; 95pp; English.
XX
CC   The present sequence is a coding sequence for a novel human abrogen,
CC   designated hATF-kringle, comprising the human urokinase plasminogen
CC   activator kringle domain. Abrogens such as hATF-kringle are potent
CC   inhibitors of endothelial proliferation and angiogenesis. Abrogen
CC   polypeptides are capable of inhibiting or reducing cell proliferation
CC   induced by both basic fibroblast growth factor (bFGF) and vascular
CC   endothelial growth factor in a specific endothelial cell proliferation
CC   assay; angiostatin only inhibits bFGF induced proliferation in this
CC   assay. Vectors that expressed abrogen polypeptides in vivo were shown to
CC   reduce tumour metastasis in 2 lung cancer models. The invention provides
CC   abrogen polypeptides and polynucleotides, and methods of using these to
CC   treat an angiogenesis-related disease or disorder, e.g. tumour metastasis
CC   (claimed)
XX
SQ   Sequence 258 BP; 61 A; 73 C; 72 G; 50 T; 0 U; 2 Other;

  Query Match             72.3%;  Score 18.8;  DB 8;  Length 258;
  Best Local Similarity   90.9%;  Pred. No. 1.3e+02;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 CTGGAAGTCTGCCAGTGTCCTT 26
              |||||| ||||||| |||||||
Db         78 CTGGAACTCTGCCACTGTCCTT 99


RESULT 9
ACC58337
ID   ACC58337 standard; cDNA; 258 BP.
XX
AC   ACC58337;
XX
DT   26-AUG-2003  (first entry)
XX
DE   Mouse abrogen (mATF-kringle) coding sequence.
XX
KW   Mouse; abrogen; kringle; mATF-kringle; angiogenesis; inhibitor;
KW   urokinase plasminogen activator; tumour; metastasis; cytostatic;
KW   gene therapy; gene; ss.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .258
FT                   /*tag=  a
FT                   /partial
FT                   /product= "Mouse abrogen"
FT                   /note= "no start or stop codon"
XX
PN   WO2003042354-A2.
XX
PD   22-MAY-2003.
XX
PF   04-SEP-2002; 2002WO-US027885.
XX
PR   04-SEP-2001; 2001US-0316300P.
XX
PA   (AVET ) AVENTIS PHARM INC.
XX
PI   Nesbit M,  Fong TC,  Brockstedt D;
XX
DR   WPI; 2003-449566/42.
DR   P-PSDB; ABR42602.
XX
PT   New abrogen polypeptide, useful for treating an angiogenesis related
PT   diseases e.g. tumor metastasis.
XX
PS   Claim 32; Page 25-26; 95pp; English.
```

```
XX
CC   The present sequence is a coding sequence for a novel murine abrogen,
CC   designated hATF-kringle, comprising the mouse urokinase plasminogen
CC   activator kringle domain. Abrogens such as mATF-kringle are potent
CC   inhibitors of endothelial proliferation and angiogenesis. Abrogen
CC   polypeptides are capable of inhibiting or reducing cell proliferation
CC   induced by both basic fibroblast growth factor (bFGF) and vascular
CC   endothelial growth factor in a specific endothelial cell proliferation
CC   assay; angiostatin only inhibits bFGF induced proliferation in this
CC   assay. Vectors that expressed abrogen polypeptides in vivo were shown to
CC   reduce tumour metastasis in 2 lung cancer models. The invention provides
CC   abrogen polypeptides and polynucleotides, and methods of using these to
CC   treat an angiogenesis-related disease or disorder, e.g. tumour metastasis
CC   (claimed)
XX
SQ   Sequence 258 BP; 62 A; 73 C; 72 G; 51 T; 0 U; 0 Other;

  Query Match             72.3%;  Score 18.8;  DB 8;  Length 258;
  Best Local Similarity   90.9%;  Pred. No. 1.3e+02;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 CTGGAAGTCTGCCAGTGTCCTT 26
              |||||| ||||||| |||||||
Db         78 CTGGAACTCTGCCACTGTCCTT 99


RESULT 10
ACC58336
ID   ACC58336 standard; cDNA; 258 BP.
XX
AC   ACC58336;
XX
DT   26-AUG-2003  (first entry)
XX
DE   Human abrogen (hATF-kringle) coding sequence.
XX
KW   Human; abrogen; kringle; hATF-kringle; angiogenesis; inhibitor;
KW   urokinase plasminogen activator; tumour; metastasis; cytostatic;
KW   gene therapy; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .258
FT                   /*tag=  a
FT                   /partial
FT                   /product= "Human abrogen"
FT                   /note= "no start or stop codon"
XX
PN   WO2003042354-A2.
XX
PD   22-MAY-2003.
XX
PF   04-SEP-2002; 2002WO-US027885.
XX
PR   04-SEP-2001; 2001US-0316300P.
XX
PA   (AVET ) AVENTIS PHARM INC.
XX
PI   Nesbit M,  Fong TC,  Brockstedt D;
XX
DR   WPI; 2003-449566/42.
DR   P-PSDB; ABR42601.
XX
PT   New abrogen polypeptide, useful for treating an angiogenesis related
PT   diseases e.g. tumor metastasis.
XX
PS   Claim 32; Page 25; 95pp; English.
XX
CC   The present sequence is a coding sequence for a novel human abrogen,
CC   designated hATF-kringle, comprising the human urokinase plasminogen
CC   activator kringle domain. Abrogens such as hATF-kringle are potent
CC   inhibitors of endothelial proliferation and angiogenesis. Abrogen
CC   polypeptides are capable of inhibiting or reducing cell proliferation
CC   induced by both basic fibroblast growth factor (bFGF) and vascular
CC   endothelial growth factor in a specific endothelial cell proliferation
CC   assay; angiostatin only inhibits bFGF induced proliferation in this
CC   assay. Vectors that expressed abrogen polypeptides in vivo were shown to
CC   reduce tumour metastasis in 2 lung cancer models. The invention provides
CC   abrogen polypeptides and polynucleotides, and methods of using these to
CC   treat an angiogenesis-related disease or disorder, e.g. tumour metastasis
CC   (claimed)
XX
SQ   Sequence 258 BP; 62 A; 74 C; 72 G; 50 T; 0 U; 0 Other;

  Query Match             72.3%;  Score 18.8;  DB 8;  Length 258;
  Best Local Similarity   90.9%;  Pred. No. 1.3e+02;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 CTGGAAGTCTGCCAGTGTCCTT 26
              |||||| ||||||| |||||||
Db         78 CTGGAACTCTGCCACTGTCCTT 99


RESULT 11
AAD27075
ID   AAD27075 standard; DNA; 264 BP.
XX
AC   AAD27075;
XX
DT   09-APR-2002  (first entry)
XX
DE   Human urokinase-type plasminogen activator (uPA) kringle DNA.
XX
KW   Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;
KW   stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;
KW   microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;
KW   tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;
KW   clotting disorder; uterine contraction disorder; respiratory disease;
KW   male impotence; adult respiratory distress syndrome; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .264
FT                   /*tag=  a
FT                   /product= "Human uPA kringle"
FT                   /note= "CDS does not include start and stop codon"
FT                   /partial
XX
PN   WO200197752-A2.
XX
PD   27-DEC-2001.
XX
PF   13-JUN-2001; 2001WO-US018976.
XX
PR   20-JUN-2000; 2000US-0212874P.
XX
PA   (UYPE-) UNIV PENNSYLVANIA.
XX
PI   Cines DB,  Higazi AA;
XX
DR   WPI; 2002-122240/16.
DR   P-PSDB; AAE16542.
XX
PT   Composition for modulating muscle cell and tissue contractility for
PT   treating atherosclerosis, asthma, hypertension, glaucoma, impotence,
PT   comprising domains from urokinase-type plasminogen activator.
XX
PS   Claim 29; Fig 1J; 117pp; English.
XX
CC   The invention relates to a composition comprising one or more domains of
CC   urokinase-type plasminogen activator (uPA). The composition is used to
CC   modulate the contractility and angiogenic activity of a mammalian muscle,
CC   endothelial cell or tissue. The composition is used for treating stroke,
CC   hypotension, hypertension, atherosclerosis, heart attack, microvascular
```

```
CC   occlusions, thrombotic microangiopathies, surgically induced thrombotic
CC   disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell
CC   invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,
CC   diabetic retinopathy, wound healing, clotting disorder, uterine
CC   contraction disorder, male impotence, respiratory disease or condition
CC   such as asthma, adult respiratory distress syndrome, primary pulmonary
CC   hypertension, microvascular thrombotic occlusion, and a disorder
CC   associated with chronic intrapulmonary fibrin formation. The present
CC   sequence is human urokinase-type plasminogen activator (uPA) kringle DNA
XX
SQ   Sequence 264 BP; 66 A; 74 C; 74 G; 50 T; 0 U; 0 Other;

  Query Match             72.3%;  Score 18.8;  DB 6;  Length 264;
  Best Local Similarity   90.9%;  Pred. No. 1.3e+02;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 CTGGAAGTCTGCCAGTGTCCTT 26
              |||||| ||||||| |||||||
Db         78 CTGGAACTCTGCCACTGTCCTT 99


RESULT 12
AAD27083
ID   AAD27083 standard; DNA; 288 BP.
XX
AC   AAD27083;
XX
DT   09-APR-2002  (first entry)
XX
DE   Human uPA kringle and connecting peptide DNA.
XX
KW   Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;
KW   stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;
KW   microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;
KW   tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;
KW   clotting disorder; uterine contraction disorder; respiratory disease;
KW   adult respiratory distress syndrome; male impotence; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .288
FT                   /*tag=  a
FT                   /product= "Human uPA kringle and connecting peptide"
FT                   /note= "CDS does not include start and stop codon"
FT                   /partial
XX
PN   WO200197752-A2.
XX
PD   27-DEC-2001.
XX
PF   13-JUN-2001; 2001WO-US018976.
XX
PR   20-JUN-2000; 2000US-0212874P.
XX
PA   (UYPE-) UNIV PENNSYLVANIA.
XX
PI   Cines DB,  Higazi AA;
XX
DR   WPI; 2002-122240/16.
DR   P-PSDB; AAE16550.
XX
PT   Composition for modulating muscle cell and tissue contractility for
PT   treating atherosclerosis, asthma, hypertension, glaucoma, impotence,
PT   comprising domains from urokinase-type plasminogen activator.
XX
PS   Claim 29; Fig 1R; 117pp; English.
XX
CC   The invention relates to a composition comprising one or more domains of
CC   urokinase-type plasminogen activator (uPA). The composition is used to
CC   modulate the contractility and angiogenic activity of a mammalian muscle,
CC   endothelial cell or tissue. The composition is used for treating stroke,
CC   hypotension, hypertension, atherosclerosis, heart attack, microvascular
CC   occlusions, thrombotic microangiopathies, surgically induced thrombotic
CC   disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell
CC   invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,
CC   diabetic retinopathy, wound healing, clotting disorder, uterine
CC   contraction disorder, male impotence, respiratory disease or condition
CC   such as asthma, adult respiratory distress syndrome, primary pulmonary
CC   hypertension, microvascular thrombotic occlusion, and a disorder
CC   associated with chronic intrapulmonary fibrin formation. The present
CC   sequence is a DNA encoding human urokinase-type plasminogen activator
CC   (uPA) kringle and connecting peptide
XX
SQ   Sequence 288 BP; 73 A; 84 C; 77 G; 54 T; 0 U; 0 Other;

  Query Match             72.3%;  Score 18.8;  DB 6;  Length 288;
  Best Local Similarity   90.9%;  Pred. No. 1.4e+02;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 CTGGAAGTCTGCCAGTGTCCTT 26
              |||||| ||||||| |||||||
Db         78 CTGGAACTCTGCCACTGTCCTT 99


RESULT 13
AAN81299
ID   AAN81299 standard; cDNA; 300 BP.
XX
AC   AAN81299;
XX
DT   25-MAR-2003  (revised)
DT   08-OCT-1990  (first entry)
XX
DE   Human prourokinase mutant designed to acquire unique restriction sites.
XX
KW   human pro-urokinase mutant; fibrinolysis; epidermal growth factor domain;
KW   ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_difference 69
FT                   /*tag=  a
FT                   /label= point mutation
FT                   /note= "Wild-type= A: Mutant= G"
FT   misc_difference 72
FT                   /*tag=  b
FT                   /label= point mutation
FT                   /note= "Wild-type= T: Mutant= C"
FT   misc_difference 209. .210
FT                   /*tag=  c
FT                   /label= two point mutations
FT                   /note= "Wild-type= GC: Mutant= CA"
XX
PN   EP253241-A.
XX
PD   20-JAN-1988.
XX
PF   03-JUL-1987;   87EP-00109628.
XX
PR   03-JUL-1986;   86JP-00156936.
PR   18-FEB-1987;   87JP-00036495.
XX
PA   (GREC ) GREEN CROSS CORP.
XX
PI   Kasai S,  Hiramatsu R,  Uno S,  Nagai M,  Arimura H;
XX
DR   WPI; 1988-015623/03.
DR   P-PSDB; AAP80996.
XX
PT   New human pro-urokinase mutants with fibrinolytic activity - have longer
PT   half-lives in blood and are obtained by recombinant DNA procedures.
XX
PS   Disclosure; Page ?; 40pp; English.
XX
```

```
CC   Site-directed mutagenesis creates two unique restriction sites; the first
CC   is recognised by SacI and the second by NdeI. The mutant sequence is
CC   inserted into an expression vector system for expression in eg E.coli.
CC   (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25-MAR-2003 to
CC   correct PI field.)
XX
SQ   Sequence 300 BP; 77 A; 82 C; 75 G; 66 T; 0 U; 0 Other;

  Query Match             72.3%;  Score 18.8;  DB 1;  Length 300;
  Best Local Similarity   90.9%;  Pred. No. 1.4e+02;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 CTGGAAGTCTGCCAGTGTCCTT 26
              |||||| ||||||| |||||||
Db        279 CTGGAACTCTGCCACTGTCCTT 300


RESULT 14
AAD27078
ID   AAD27078 standard; DNA; 405 BP.
XX
AC   AAD27078;
XX
DT   09-APR-2002  (first entry)
XX
DE   Human urokinase-type plasminogen activator amino terminal fragment DNA.
XX
KW   Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;
KW   stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;
KW   microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;
KW   tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;
KW   clotting disorder; uterine contraction disorder; respiratory disease;
KW   adult respiratory distress syndrome; amino terminal fragment; ATF;
KW   male impotence; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .405
FT                   /*tag=  a
FT                   /product= "Human uPA ATF"
FT                   /note= "CDS does not include start and stop codon"
FT                   /partial
XX
PN   WO200197752-A2.
XX
PD   27-DEC-2001.
XX
PF   13-JUN-2001; 2001WO-US018976.
XX
PR   20-JUN-2000; 2000US-0212874P.
XX
PA   (UYPE-) UNIV PENNSYLVANIA.
XX
PI   Cines DB,  Higazi AA;
XX
DR   WPI; 2002-122240/16.
DR   P-PSDB; AAE16545.
XX
PT   Composition for modulating muscle cell and tissue contractility for
PT   treating atherosclerosis, asthma, hypertension, glaucoma, impotence,
PT   comprising domains from urokinase-type plasminogen activator.
XX
PS   Claim 29; Fig 1M; 117pp; English.
XX
CC   The invention relates to a composition comprising one or more domains of
CC   urokinase-type plasminogen activator (uPA). The composition is used to
CC   modulate the contractility and angiogenic activity of a mammalian muscle,
CC   endothelial cell or tissue. The composition is used for treating stroke,
CC   hypotension, hypertension, atherosclerosis, heart attack, microvascular
CC   occlusions, thrombotic microangiopathies, surgically induced thrombotic
CC   disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell
CC   invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,
CC   diabetic retinopathy, wound healing, clotting disorder, uterine
CC   contraction disorder, male impotence, respiratory disease or condition
CC   such as asthma, adult respiratory distress syndrome, primary pulmonary
CC   hypertension, microvascular thrombotic occlusion, and a disorder
CC   associated with chronic intrapulmonary fibrin formation. The present
CC   sequence is human urokinase-type plasminogen activator (uPA) amino
CC   terminal fragment (ATF) DNA
XX
SQ   Sequence 405 BP; 113 A; 105 C; 104 G; 83 T; 0 U; 0 Other;

  Query Match             72.3%;  Score 18.8;  DB 6;  Length 405;
  Best Local Similarity   90.9%;  Pred. No. 1.4e+02;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 CTGGAAGTCTGCCAGTGTCCTT 26
              |||||| ||||||| |||||||
Db        219 CTGGAACTCTGCCACTGTCCTT 240


RESULT 15
AAD27082
ID   AAD27082 standard; DNA; 429 BP.
XX
AC   AAD27082;
XX
DT   09-APR-2002  (first entry)
XX
DE   Human uPA amino terminal fragment (ATF) and connecting peptide DNA.
XX
KW   Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;
KW   stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;
KW   microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;
KW   tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;
KW   clotting disorder; uterine contraction disorder; respiratory disease;
KW   adult respiratory distress syndrome; amino terminal fragment; ATF;
KW   male impotence; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .429
FT                   /*tag=  a
FT                   /product= "Human uPA ATF and connecting peptide"
FT                   /note= "CDS does not include start and stop codon"
FT                   /partial
XX
PN   WO200197752-A2.
XX
PD   27-DEC-2001.
XX
PF   13-JUN-2001; 2001WO-US018976.
XX
PR   20-JUN-2000; 2000US-0212874P.
XX
PA   (UYPE-) UNIV PENNSYLVANIA.
XX
PI   Cines DB,  Higazi AA;
XX
DR   WPI; 2002-122240/16.
DR   P-PSDB; AAE16549.
XX
PT   Composition for modulating muscle cell and tissue contractility for
PT   treating atherosclerosis, asthma, hypertension, glaucoma, impotence,
PT   comprising domains from urokinase-type plasminogen activator.
XX
PS   Claim 29; Fig 1Q; 117pp; English.
XX
CC   The invention relates to a composition comprising one or more domains of
CC   urokinase-type plasminogen activator (uPA). The composition is used to
CC   modulate the contractility and angiogenic activity of a mammalian muscle,
CC   endothelial cell or tissue. The composition is used for treating stroke,
CC   hypotension, hypertension, atherosclerosis, heart attack, microvascular
CC   occlusions, thrombotic microangiopathies, surgically induced thrombotic
```

CC disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell
CC invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,
CC diabetic retinopathy, wound healing, clotting disorder, uterine
CC contraction disorder, male impotence, respiratory disease or condition
CC such as asthma, adult respiratory distress syndrome, primary pulmonary
CC hypertension, microvascular thrombotic occlusion, and a disorder
CC associated with chronic intrapulmonary fibrin formation. The present
CC sequence is a DNA encoding human urokinase-type plasminogen activator
CC (uPA) amino terminal fragment (ATF) and connecting peptide
XX
SQ Sequence 429 BP; 120 A; 115 C; 107 G; 87 T; 0 U; 0 Other;

Query Match 72.3%; Score 18.8; DB 6; Length 429;
Best Local Similarity 90.9%; Pred. No. 1.4e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          5 CTGGAAGTCTGCCAGTGTCCTT 26
              |||||| ||||||| |||||||
Db        219 CTGGAACTCTGCCACTGTCCTT 240
```

Search completed: February 19, 2004, 21:51:43
Job time : 25.0879 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 21:21:54 ; Search time 172.149 Seconds
(without alignments)
4510.152 Million cell updates/sec

Title: US-09-939-853A-141
Perfect score: 26
Sequence: 1 ccttctggaagtctgccagtgtcctt 26

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	26	100.0	616	13	BX383606	BX383606 BX383606
c	2	26	100.0	778	12	BG178487	BG178487 602328305
c	3	26	100.0	878	12	BQ053486	BQ053486 AGENCOURT
c	4	26	100.0	986	12	BQ054265	BQ054265 AGENCOURT
c	5	26	100.0	1020	12	BQ054281	BQ054281 AGENCOURT
c	6	26	100.0	1201	9	AL541041	AL541041 AL541041
c	7	21.8	83.8	354	12	BM483329	BM483329 536609 MA
	8	21.8	83.8	389	28	AZ620995	AZ620995 1M0453M21
c	9	21.8	83.8	541	12	BI898999	BI898999 480839 MA
c	10	20.2	77.7	442	12	BG228158	BG228158 mac35e07.
c	11	20.2	77.7	689	10	BB468570	BB468570 BB468570
c	12	20.2	77.7	693	10	BB195256	BB195256 BB195256
	13	20.2	77.7	1226	11	AK011064	AK011064 Mus muscu
c	14	20.2	77.7	3016	11	AK039689	AK039689 Mus muscu
c	15	19.8	76.2	276	29	CG508158	CG508158 OST58614
c	16	19.8	76.2	518	9	AA747812	AA747812 nx86c07.s
c	17	19.6	75.4	302	29	CG635108	CG635108 OST356941
c	18	19.6	75.4	389	10	BE015229	BE015229 127457 MA
c	19	19.6	75.4	439	10	BE015226	BE015226 127451 MA
c	20	19.6	75.4	447	14	CA893892	CA893892 B0182A06-
	21	19.6	75.4	533	9	AI551288	AI551288 vo72d04.x
c	22	19.6	75.4	578	14	CA894739	CA894739 B0187E08-
c	23	19.6	75.4	967	12	BI192570	BI192570 602945473
c	24	19.6	75.4	1114	28	CC250655	CC250655 CH261-75M
c	25	19.6	75.4	1132	28	CC286280	CC286280 CH261-100
c	26	19.6	75.4	1281	28	CC283225	CC283225 CH261-26L
	27	19.4	74.6	546	13	C78860	C78860 C78860 Mous
c	28	19.2	73.8	129	29	CG521939	CG521939 OST90738
c	29	19.2	73.8	206	29	CG558956	CG558956 OST177066
c	30	19.2	73.8	208	29	CG555648	CG555648 OST170490
c	31	19.2	73.8	213	29	CG576620	CG576620 OST211551
c	32	19.2	73.8	249	29	CG639953	CG639953 OST372089
c	33	19.2	73.8	284	29	CG648011	CG648011 OST399304
c	34	19.2	73.8	287	29	CG618235	CG618235 OST312776
c	35	19.2	73.8	287	29	CG630122	CG630122 OST345077
c	36	19.2	73.8	289	29	CG578637	CG578637 OST216073
c	37	19.2	73.8	291	29	CG645423	CG645423 OST389989
	38	19.2	73.8	348	9	AI163788	AI163788 A048p58u
c	39	19.2	73.8	403	14	T74108	T74108 yc55f12.s1
c	40	19.2	73.8	438	14	T61390	T61390 yb90d07.s1
	41	19.2	73.8	521	29	CE370901	CE370901 tigr-gss-
c	42	19.2	73.8	553	9	AA991151	AA991151 os40a08.s
c	43	19.2	73.8	606	28	BH310531	BH310531 CH230-101
c	44	19.2	73.8	661	29	AG043855	AG043855 Pan trogl
c	45	19.2	73.8	749	28	BZ204700	BZ204700 CH230-487

ALIGNMENTS

RESULT 1
BX383606/c
LOCUS BX383606 616 bp mRNA linear EST 08-MAY-2003
DEFINITION BX383606 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
Homo sapiens cDNA clone CS0DJ013YK10 5-PRIME, mRNA sequence.
ACCESSION BX383606
VERSION BX383606.1 GI:30457152
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 616)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 9825.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DJ013BF05QP1&cluster=9825.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ InVitroGen Corporation 1600

```
            Faraday Avenue Genoscope sequence ID : CS0DJ013BF05QP1.
FEATURES             Location/Qualifiers
     source          1. .616
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DJ013YK10"
                     /cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
                     /cell_line="JURKAT"
                     /clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
                     10-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

  Query Match             100.0%;  Score 26;  DB 13;  Length 616;
  Best Local Similarity   100.0%;  Pred. No. 1.4;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCTTCTGGAAGTCTGCCAGTGTCCTT 26
              ||||||||||||||||||||||||||
Db        325 CCTTCTGGAAGTCTGCCAGTGTCCTT 300


RESULT 2
BG178487/c
LOCUS       BG178487                778 bp    mRNA    linear   EST 06-FEB-2001
DEFINITION  602328305F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4429896 5',
            mRNA sequence.
ACCESSION   BG178487
VERSION     BG178487.1  GI:12685190
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 778)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10182  row: i  column: 01
            High quality sequence stop: 657.
FEATURES             Location/Qualifiers
     source          1. .778
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4429896"
                     /tissue_type="adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_91"
                     /note="Organ: prostate; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.4 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             100.0%;  Score 26;  DB 12;  Length 778;
  Best Local Similarity   100.0%;  Pred. No. 1.5;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCTTCTGGAAGTCTGCCAGTGTCCTT 26
              ||||||||||||||||||||||||||
Db        240 CCTTCTGGAAGTCTGCCAGTGTCCTT 215


RESULT 3
BQ053486/c
LOCUS       BQ053486                878 bp    mRNA    linear   EST 29-MAR-2002
DEFINITION  AGENCOURT_6822017 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5935253
            5', mRNA sequence.
ACCESSION   BQ053486
VERSION     BQ053486.1  GI:19812826
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 878)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2122  row: l  column: 06
            High quality sequence stop: 394.
FEATURES             Location/Qualifiers
     source          1. .878
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5935253"
                     /tissue_type="natural killer cells, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_106"
                     /note="Organ: blood; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Library constructed  by Ling Hong in the
                     laboratory of Gerald M. Rubin (University of  California,
                     Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                     Superscript II RT (Life Technologies). Note: this is a
                     NIH_MGC Library."
ORIGIN

  Query Match             100.0%;  Score 26;  DB 12;  Length 878;
  Best Local Similarity   100.0%;  Pred. No. 1.5;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCTTCTGGAAGTCTGCCAGTGTCCTT 26
              ||||||||||||||||||||||||||
Db        151 CCTTCTGGAAGTCTGCCAGTGTCCTT 126


RESULT 4
BQ054265/c
LOCUS       BQ054265                986 bp    mRNA    linear   EST 29-MAR-2002
DEFINITION  AGENCOURT_6830248 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5936339
            5', mRNA sequence.
ACCESSION   BQ054265
VERSION     BQ054265.1  GI:19813605
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 986)
```

```
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2125  row: i  column: 12
            High quality sequence stop: 515.
FEATURES             Location/Qualifiers
     source          1. .986
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5936339"
                     /tissue_type="natural killer cells, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_106"
                     /note="Organ: blood; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Library constructed  by Ling Hong in the
                     laboratory of Gerald M. Rubin (University of  California,
                     Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                     Superscript II RT (Life Technologies). Note: this is a
                     NIH_MGC Library."
ORIGIN

  Query Match             100.0%;  Score 26;  DB 12;  Length 986;
  Best Local Similarity   100.0%;  Pred. No. 1.6;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCTTCTGGAAGTCTGCCAGTGTCCTT 26
              ||||||||||||||||||||||||||
Db        157 CCTTCTGGAAGTCTGCCAGTGTCCTT 132


RESULT 5
BQ054281/c
LOCUS       BQ054281                1020 bp    mRNA    linear   EST 29-MAR-2002
DEFINITION  AGENCOURT_6830234 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5936362
            5', mRNA sequence.
ACCESSION   BQ054281
VERSION     BQ054281.1  GI:19813621
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1020)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2125  row: j  column: 11
            High quality sequence stop: 556.
FEATURES             Location/Qualifiers
     source          1. .1020
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5936362"
                     /tissue_type="natural killer cells, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_106"
                     /note="Organ: blood; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Library constructed  by Ling Hong in the
                     laboratory of Gerald M. Rubin (University of  California,
                     Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                     Superscript II RT (Life Technologies). Note: this is a
                     NIH_MGC Library."
ORIGIN

  Query Match             100.0%;  Score 26;  DB 12;  Length 1020;
  Best Local Similarity   100.0%;  Pred. No. 1.6;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCTTCTGGAAGTCTGCCAGTGTCCTT 26
              ||||||||||||||||||||||||||
Db         90 CCTTCTGGAAGTCTGCCAGTGTCCTT 65


RESULT 6
AL541041/c
LOCUS       AL541041                1201 bp    mRNA    linear   EST 12-MAY-2003
DEFINITION  AL541041 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE005YK23
            5-PRIME, mRNA sequence.
ACCESSION   AL541041
VERSION     AL541041.2  GI:30544829
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1201)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 15, 2001 this sequence version replaced gi:12871733.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 9825.r For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DE005AF12QP1&cluster=9825.r. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DE005AF12QP1.
FEATURES             Location/Qualifiers
     source          1. .1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DE005YK23"
                     /tissue_type="PLACENTA"
                     /clone_lib="Homo sapiens PLACENTA"
                     /note="Vector: pCMVSPORT_6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."
ORIGIN

  Query Match             100.0%;  Score 26;  DB 9;  Length 1201;
  Best Local Similarity   100.0%;  Pred. No. 1.7;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCTTCTGGAAGTCTGCCAGTGTCCTT 26
```

```
                    ||||||||||||||||||||||||
Db        337 CCTTCTGGAAGTCTGCCAGTGTCCTT 312


RESULT 7
BM483329/c
LOCUS       BM483329                354 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  536609 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION   BM483329
VERSION     BM483329.1  GI:18533657
KEYWORDS    EST.
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 354)
  AUTHORS   Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
            Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C.,
            Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A.,
            Chitko-McKown,C.G., Pertea,G., Holt,I., Karamycheva,S., Liang,F.,
            Quackenbush,J. and Keele,J.W.
  TITLE     Sequence evaluation of four pooled-tissue normalized bovine cDNA
            libraries and construction of a gene index for cattle
  JOURNAL   Genome Res. 11 (4), 626-630 (2001)
  MEDLINE   21180013
   PUBMED   11282978
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and alt_trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -minmatch 12 options.
            PCR PRimers
            FORWARD: AGGAAACAGCTATGACCAT
            BACKWARD: GTTTTCCCAGTCACGACG
            Plate: 6  row: C  column: 14
            Seq primer: ATTTAGGTGACACTATAG.
FEATURES             Location/Qualifiers
     source          1. .354
                     /organism="Bos taurus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9913"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /clone_lib="MARC 3BOV"
                     /note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
                     Library made from pooled tissue from marrow, alveolar
                     macrophage, ovary, fetal semitendonosus muscle, and fetal
                     longissimus muscle."
ORIGIN

  Query Match             83.8%;  Score 21.8;  DB 12;  Length 354;
  Best Local Similarity   92.0%;  Pred. No. 72;
  Matches   23;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 CTTCTGGAAGTCTGCCAGTGTCCTT 26
              ||||| |||||||||||| |||||||
Db        245 CTTCTAGAAGTCTGCCAATGTCCTT 221


RESULT 8
AZ620995
LOCUS       AZ620995                389 bp    DNA     linear   GSS 13-DEC-2000
DEFINITION  1M0453M21R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0453M21 R, genomic survey sequence.
ACCESSION   AZ620995
VERSION     AZ620995.1  GI:11743185
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 389)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0453  row: M  column: 21
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 389.
FEATURES             Location/Qualifiers
     source          1. .389
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0453M21"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
ORIGIN

  Query Match             83.8%;  Score 21.8;  DB 28;  Length 389;
  Best Local Similarity   92.0%;  Pred. No. 74;
  Matches   23;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 CTTCTGGAAGTCTGCCAGTGTCCTT 26
              ||||||||| |||||||||||| ||
Db         80 CTTCTGGAATTCTGCCAGTGTCTTT 104


RESULT 9
BI898999/c
LOCUS       BI898999                541 bp    mRNA    linear   EST 16-OCT-2001
DEFINITION  480839 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION   BI898999
VERSION     BI898999.1  GI:16187055
KEYWORDS    EST.
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 541)
  AUTHORS   Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
            Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C.,
            Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A.,
            Chitko-McKown,C.G., Pertea,G., Holt,I., Karamycheva,S., Liang,F.,
            Quackenbush,J. and Keele,J.W.
  TITLE     Sequence evaluation of four pooled-tissue normalized bovine cDNA
            libraries and construction of a gene index for cattle
  JOURNAL   Genome Res. 11 (4), 626-630 (2001)
  MEDLINE   21180013
   PUBMED   11282978
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and alt_trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -minmatch 12 options.
            PCR PRimers
            FORWARD: AGGAAACAGCTATGACCAT
            BACKWARD: GTTTTCCCAGTCACGACG
            Plate: 105  row: P  column: 20
            Seq primer: ATTTAGGTGACACTATAG.
FEATURES             Location/Qualifiers
     source          1. .541
                     /organism="Bos taurus"
                     /mol_type="mRNA"
                     /db_xref="taxon:9913"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /clone_lib="MARC 2BOV"
                     /note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
                     Library made from pooled tissue from testis, thymus,
                     semitendonosus muscle, longissimus muscle, pancreas,
                     adrenal, and endometrium."
ORIGIN

  Query Match             83.8%;  Score 21.8;  DB 12;  Length 541;
  Best Local Similarity   92.0%;  Pred. No. 83;
  Matches   23;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 CTTCTGGAAGTCTGCCAGTGTCCTT 26
              ||||| |||||||||||| |||||||
Db        198 CTTCTAGAAGTCTGCCAATGTCCTT 174


RESULT 10
BG228158/c
LOCUS       BG228158                442 bp    mRNA    linear   EST 08-FEB-2001
DEFINITION  mac35e07.x1 Soares mouse 3NbMS Mus musculus cDNA clone
            IMAGE:4001772 3', mRNA sequence.
ACCESSION   BG228158
VERSION     BG228158.1  GI:12715662
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 442)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:1499500
            High quality sequence stop: 402.
FEATURES             Location/Qualifiers
     source          1. .442
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4001772"
                     /sex="male"
                     /tissue_type="Spleen"
                     /dev_stage="4 weeks"
                     /lab_host="DH10B"
                     /clone_lib="Soares mouse 3NbMS"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGCTGTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. RNA
                     provided by Dr. Bertrand Jordan. Library went through
                     three rounds of normalization, and was constructed by
                     Bento Soares and M.Fatima Bonaldo."
ORIGIN

  Query Match             77.7%;  Score 20.2;  DB 12;  Length 442;
  Best Local Similarity   88.0%;  Pred. No. 3.8e+02;
  Matches   22;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 CTTCTGGAAGTCTGCCAGTGTCCTT 26
              ||||||||| |||| || |||||||
Db        232 CTTCTGGAACTCTGACAATGTCCTT 208


RESULT 11
BB468570/c
LOCUS       BB468570                689 bp    mRNA    linear   EST 25-OCT-2001
DEFINITION  BB468570 RIKEN full-length enriched, 12 days embryo eyeball Mus
            musculus cDNA clone D230019P08 3', mRNA sequence.
ACCESSION   BB468570
VERSION     BB468570.2  GI:16427534
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 689)
  AUTHORS   Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,
            Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,
            Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,
            Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,
            Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
            Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
  JOURNAL   Unpublished (2001)
COMMENT     On Jul 22, 2000 this sequence version replaced gi:9385759.
            Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
```

Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. . 10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and Hayashizaki,Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

FEATURES Location/Qualifiers
source 1. .689
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="D230019P08"
/tissue_type="eyeball"
/dev_stage="12 days embryo"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 12 days embryo eyeball"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAGCGGCCGCAACTCGAGTTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

ORIGIN

Query Match 77.7%; Score 20.2; DB 10; Length 689;
Best Local Similarity 88.0%; Pred. No. 4.4e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 CTTCTGGAAGTCTGCCAGTGTCCTT 26
||||||||||.|.||||||||.|||
Db 95 CTTCTGGAAGACAGCCAGTGTTCTT 71

RESULT 12
BB195256/c
LOCUS BB195256 693 bp mRNA linear EST 19-OCT-2001
DEFINITION BB195256 RIKEN full-length enriched, adult male spinal cord Mus musculus cDNA clone A330087I24 3', mRNA sequence.
ACCESSION BB195256
VERSION BB195256.2 GI:16270936
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 693)
AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
TITLE RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
JOURNAL Unpublished (2001)
COMMENT On Jun 30, 2000 this sequence version replaced gi:8855875.
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. . 10 (10), 1617-1630 (2000)
wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. . 10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and Hayashizaki,Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

FEATURES Location/Qualifiers
source 1. .693
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="A330087I24"
/sex="male"
/tissue_type="spinal cord"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, adult male spinal cord"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC

3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. "
ORIGIN

Query Match 77.7%; Score 20.2; DB 10; Length 693;
Best Local Similarity 88.0%; Pred. No. 4.5e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 CTTCTGGAAGTCTGCCAGTGTCCTT 26
||||||||| |||| || |||||||
Db 595 CTTCTGGAACTCTGACAATGTCCTT 571

RESULT 13
AK011064
LOCUS AK011064 1226 bp mRNA linear HTC 20-SEP-2003
DEFINITION Mus musculus 13 days embryo liver cDNA, RIKEN full-length enriched library, clone:2510040D07 product:hypothetical Signal peptidase/Trp-Asp (WD) repeats profile/Trp-Asp (WD) repeats circular profile/G-protein beta WD-40 repeats containing protein, full insert sequence.
ACCESSION AK011064
VERSION AK011064.1 GI:12846940
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 1226)
AUTHORS Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOLR.
FEATURES Location/Qualifiers
source 1. .1226
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:2510040D07"
/db_xref="MGI:1902545"
/db_xref="taxon:10090"
/clone="2510040D07"
/tissue_type="liver"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="13 days embryo"
CDS 60. .1130
/note="unnamed protein product; hypothetical Signal peptidase/Trp-Asp (WD) repeats profile/Trp-Asp (WD) repeats circular profile/G-protein beta WD-40 repeats containing protein (InterPro|IPR000508, PROSITE|PS50082, PROSITE|PS50294, InterPro|IPR001680, evidence: InterPro) putative"
/codon_start=1
/protein_id="BAB27371.1"
/db_xref="GI:12846941"
/translation="MAAPSQDPSLERHFKGHRDAVTRVDFSLNTKHLASGSMDSTLMI WHMKPQSRAYRFTGHNDAVTCVNFSPSGHLLASGSRDKTVRIWVPNVKGESTVFRAHT ATVRSVHFCSDGQSLVTASDDKTVKVWSTHRQRFLFSLTQHINWVRCAKFSPDGRLIV SASDDKTVKLWDKTSRECIHSYCEHGGFVTYVDFHPSGTCIAAAGMDNTVKVWDARTH RLLQHYQLHSAAVNALSFHPSGNYLITASSDSTLKILDLMEGRLLYTLHGHQGPATTV AFSRTGEYFASGGSDEQVMVWKSNFDIVDYGDMKARRPPPLTSSSGTLTVSILEQRLT LTEDRLKQCLENQQLIMQRTPP"
polyA_signal 1208. .1213
/note="putative"
polyA_site 1226
/note="putative"
ORIGIN

Query Match 77.7%; Score 20.2; DB 11; Length 1226;
Best Local Similarity 88.0%; Pred. No. 5.5e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 CTTCTGGAAGTCTGCCAGTGTCCTT 26
||||||||| |||| |||||||| |
Db 1018 CTTCTGGAACTCTGACAGTGTCCAT 1042

```
RESULT 14
AK039689/c
LOCUS       AK039689               3016 bp    mRNA    linear   HTC 19-SEP-2003
DEFINITION  Mus musculus adult male spinal cord cDNA, RIKEN full-length
            enriched library, clone:A330087I24 product:unclassifiable, full
            insert sequence.
ACCESSION   AK039689
VERSION     AK039689.1  GI:26087341
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5
  AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6  (bases 1 to 3016)
  AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
            Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
            Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
            Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
            Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
            Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
            Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
            Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
            Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
            Please visit our web site for further details.
            URL:http://genome.gsc.riken.go.jp/
            URL:http://fantom.gsc.riken.go.jp/.
FEATURES             Location/Qualifiers
     source          1. .3016
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:A330087I24"
                     /db_xref="MGI:2404195"
                     /db_xref="taxon:10090"
                     /clone="A330087I24"
                     /sex="male"
                     /tissue_type="spinal cord"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"
     misc_feature    1. .3016
                     /note="unclassifiable"
ORIGIN

  Query Match             77.7%;  Score 20.2;  DB 11;  Length 3016;
  Best Local Similarity   88.0%;  Pred. No. 7.6e+02;
  Matches   22;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 CTTCTGGAAGTCTGCCAGTGTCCTT 26
              ||||||||| |||| || |||||||
Db       2918 CTTCTGGAACTCTGACAATGTCCTT 2894


RESULT 15
CG508158/c
LOCUS       CG508158                276 bp    DNA     linear   GSS 01-OCT-2003
DEFINITION  OST58614 Mus musculus 129Sv/Ev Mus musculus genomic clone OST58614,
            genomic survey sequence.
ACCESSION   CG508158
VERSION     CG508158.1  GI:37290088
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 276)
  AUTHORS   Zambrowicz,B.P., Abuin,A., Ramirez-Solis,R., Richter,L.J.,
            Piggott,J., BeltrandelRio,H., Buxton,E.C., Edwards,J., Finch,R.A.,
            Friddle,C.J., Gupta,A., Hansen,G., Hu,Y., Huang,W., Jaing,C.,
            Key,B.W. Jr., Kipp,P., Kohlhauff,B., Ma,Z.-Q., Markesich,D.,
            Payne,R., Potter,D.G., Qian,N., Shaw,J., Schrick,J., Shi,Z.-Z.,
            Sparks,M.J., Van Sligtenhorst,I., Vogel,P., Walke,W., Xu,N.,
            Zhu,Q., Person,C. and Sands,A.T.
  TITLE     Wnk1 kinase deficiency lowers blood pressure in mice: a gene-trap
            screen to identify potential targets for therapeutic intervention
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)
COMMENT     Contact: Zambrowicz BP
            OmniBank
            Lexicon Genetics Incorporated
            4000 Research Forest Drive, The Woodlands, TX 77381, USA
            Email: materials@lexgen.com
            Gene trap sequence tag generated by 3' RACE from mouse ES cells as
            described in Zambrowicz et al (Nature. 1998 Apr 9;392(6676):608-11)
            Class: Gene Trap.
FEATURES             Location/Qualifiers
     source          1. .276
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="129SV/EV"
                     /db_xref="taxon:10090"
                     /clone="OST58614"
                     /cell_type="embryonic stem cell"
                     /clone_lib="Mus musculus 129Sv/Ev"
```

ORIGIN

Query Match 76.2%; Score 19.8; DB 29; Length 276;
Best Local Similarity 84.0%; Pred. No. 4.8e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy          2 CTTCTGGAAGTCTGCCAGTGTCCTT 26
              ||||||||| | || || |||||||
Db        183 CTTCTGGAACTNTGNCAATGTCCTT 159
```

Search completed: February 20, 2004, 01:41:40
Job time : 176.149 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 21:32:00 ; Search time 4.21902 Seconds
(without alignments)
3419.919 Million cell updates/sec

Title: US-09-939-853A-141
Perfect score: 26
Sequence: 1 ccttctggaagtctgccagtgtcctt 26

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	18.8	72.3	423	2	US-08-797-689-11	Sequence 11, Appl
	2	18.8	72.3	423	4	US-09-984-186-11	Sequence 11, Appl
	3	18.8	72.3	600	4	US-09-101-272G-72	Sequence 72, Appl
	4	18.8	72.3	624	4	US-09-101-272G-79	Sequence 79, Appl
	5	18.8	72.3	645	4	US-09-101-272G-95	Sequence 95, Appl
	6	18.8	72.3	666	4	US-09-101-272G-97	Sequence 97, Appl
	7	18.8	72.3	1233	1	US-08-254-922-1	Sequence 1, Appli
	8	18.8	72.3	1233	1	US-08-286-748B-1	Sequence 1, Appli
	9	18.8	72.3	1236	1	US-07-957-039A-7	Sequence 7, Appli
	10	18.8	72.3	1236	1	US-08-153-799-17	Sequence 17, Appl
	11	18.8	72.3	1236	4	US-09-023-655-927	Sequence 927, App
	12	18.8	72.3	1372	6	5219569-1	Patent No. 5219569
	13	18.8	72.3	1475	4	US-09-643-597-122	Sequence 122, App
	14	18.8	72.3	1475	4	US-09-480-884A-122	Sequence 122, App
	15	18.8	72.3	1475	4	US-09-542-615A-122	Sequence 122, App
	16	18.8	72.3	1475	4	US-09-606-421B-122	Sequence 122, App
	17	18.8	72.3	1475	4	US-09-221-107-122	Sequence 122, App
	18	18.8	72.3	2294	4	US-09-643-597-123	Sequence 123, App
	19	18.8	72.3	2294	4	US-09-480-884A-123	Sequence 123, App
	20	18.8	72.3	2294	4	US-09-542-615A-123	Sequence 123, App
	21	18.8	72.3	2294	4	US-09-606-421B-123	Sequence 123, App
	22	18.8	72.3	2294	4	US-09-023-655-1217	Sequence 1217, Ap
	23	18.8	72.3	2294	4	US-09-221-107-123	Sequence 123, App
	24	18.8	72.3	2301	6	5188829-2	Patent No. 5188829
	25	17.8	68.5	9391	4	US-09-562-702A-11	Sequence 11, Appl
	26	17.8	68.5	9511	4	US-09-562-702A-9	Sequence 9, Appli
c	27	17.6	67.7	273	4	US-09-313-294A-38	Sequence 38, Appl
c	28	17.6	67.7	908	4	US-09-800-729-25	Sequence 25, Appl
	29	17.2	66.2	1335	1	US-07-942-157A-2	Sequence 2, Appli
c	30	17	65.4	274	4	US-09-313-294A-3335	Sequence 3335, Ap
c	31	17	65.4	449	4	US-09-621-976-14601	Sequence 14601, A
c	32	17	65.4	501	4	US-09-621-976-1584	Sequence 1584, Ap
	33	17	65.4	1001	4	US-09-641-638-285	Sequence 285, App
c	34	17	65.4	3364	2	US-08-735-609-9	Sequence 9, Appli
c	35	17	65.4	3364	2	US-08-735-609-9	Sequence 9, Appli
c	36	17	65.4	3364	3	US-09-315-372-9	Sequence 9, Appli
c	37	17	65.4	3364	3	US-09-244-752-9	Sequence 9, Appli
c	38	17	65.4	3364	3	US-09-245-497-9	Sequence 9, Appli
c	39	17	65.4	3364	4	US-09-562-919-9	Sequence 9, Appli
c	40	17	65.4	8147	4	US-09-514-247A-9	Sequence 9, Appli
c	41	16.6	63.8	2157	1	US-08-336-618-25	Sequence 25, Appl
c	42	16.6	63.8	2200	2	US-08-462-481-3	Sequence 3, Appli
c	43	16.6	63.8	2200	2	US-08-436-771-5	Sequence 5, Appli
c	44	16.6	63.8	2200	2	US-08-434-998-5	Sequence 5, Appli
c	45	16.6	63.8	2200	2	US-08-487-797-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-08-797-689-11
; Sequence 11, Application US/08797689
; Patent No. 5876969
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fournier, Alain
; APPLICANT: Guitton, Jean-Dominique
; APPLICANT: Jung, Gerard
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; TITLE OF INVENTION: PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 423 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double

```
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  3..419
US-08-797-689-11

  Query Match             72.3%;  Score 18.8;  DB 2;  Length 423;
  Best Local Similarity   90.9%;  Pred. No. 11;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 CTGGAAGTCTGCCAGTGTCCTT 26
              |||||| ||||||| |||||||
Db        230 CTGGAACTCTGCCACTGTCCTT 251


RESULT 2
US-09-984-186-11
; Sequence 11, Application US/09984186
; Patent No. 6686179
;   GENERAL INFORMATION:
;        APPLICANT: Fleer, Reinhard
;                   Fournier, Alain
;                   Guitton, Jean-Dominique
;                   Jung, Gerard
;                   Yeh, Patrice
;        TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
;                            PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
;                            CONTAINING SAID POLYPEPTIDES
;        NUMBER OF SEQUENCES: 36
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Rhone-Poulenc Rorer Inc.
;             STREET: 500 Arcola Road, 3C43
;             CITY: Collegeville
;             STATE: PA
;             COUNTRY: USA
;             ZIP: 19426
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: Macintosh
;             OPERATING SYSTEM: System 7.1
;             SOFTWARE: Word 5.1 (PatentIn)
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/984,186
;             FILING DATE: 29-Oct-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/08/797,689
;             FILING DATE: 31-JAN-1997
;             APPLICATION NUMBER: US 08/256,927
;             FILING DATE: 28-JUL-1994
;             APPLICATION NUMBER: FR 92/01064
;             FILING DATE: 31-JAN-1992
;             APPLICATION NUMBER: PCT/FR93/00085
;             FILING DATE: 28-JAN-1993
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Smith Ph.D., Julie K.
;             REGISTRATION NUMBER: P-38,619
;             REFERENCE/DOCKET NUMBER: ST92006-US
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (610) 454-3839
;             TELEFAX: (610) 454-3808
;   INFORMATION FOR SEQ ID NO: 11:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 423 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        FEATURE:
;             NAME/KEY:  CDS
;             LOCATION:  3..419
;        SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-09-984-186-11

  Query Match             72.3%;  Score 18.8;  DB 4;  Length 423;
  Best Local Similarity   90.9%;  Pred. No. 11;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 CTGGAAGTCTGCCAGTGTCCTT 26
              |||||| ||||||| |||||||
Db        230 CTGGAACTCTGCCACTGTCCTT 251


RESULT 3
US-09-101-272G-72
; Sequence 72, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
;  APPLICANT: Nissin Food Products Co., Ltd.
;  TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
;  FILE REFERENCE: Q50979
;  CURRENT APPLICATION NUMBER: US/09/101,272G
;  CURRENT FILING DATE:  1998-07-08
;  PRIOR APPLICATION NUMBER: JP 1059/1996
;  PRIOR FILING DATE: 1996-01-08
;  NUMBER OF SEQ ID NOS: 107
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 72
;   LENGTH: 600
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: ATF domain of uPA
;   NAME/KEY: CDS
;   LOCATION: (1)..(600)
;   OTHER INFORMATION:
;   NAME/KEY: mat_peptide
;   LOCATION: (61)..()
;   OTHER INFORMATION:
US-09-101-272G-72

  Query Match             72.3%;  Score 18.8;  DB 4;  Length 600;
  Best Local Similarity   90.9%;  Pred. No. 12;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 CTGGAAGTCTGCCAGTGTCCTT 26
              |||||| ||||||| |||||||
Db        279 CTGGAACTCTGCCACTGTCCTT 300


RESULT 4
US-09-101-272G-79
; Sequence 79, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
;  APPLICANT: Nissin Food Products Co., Ltd.
;  TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
;  FILE REFERENCE: Q50979
;  CURRENT APPLICATION NUMBER: US/09/101,272G
;  CURRENT FILING DATE:  1998-07-08
;  PRIOR APPLICATION NUMBER: JP 1059/1996
;  PRIOR FILING DATE: 1996-01-08
;  NUMBER OF SEQ ID NOS: 107
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 79
;   LENGTH: 624
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: ATFHI chimeric protein
;   NAME/KEY: CDS
;   LOCATION: (12)..(593)
;   OTHER INFORMATION:
;   NAME/KEY: mat_peptide
```

```
;   LOCATION: (15)..()
;   OTHER INFORMATION:
US-09-101-272G-79

  Query Match             72.3%;  Score 18.8;  DB 4;  Length 624;
  Best Local Similarity   90.9%;  Pred. No. 12;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 CTGGAAGTCTGCCAGTGTCCTT 26
              |||||| ||||||| |||||||
Db        233 CTGGAACTCTGCCACTGTCCTT 254


RESULT 5
US-09-101-272G-95
; Sequence 95, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
;  APPLICANT: Nissin Food Products Co., Ltd.
;  TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
;  FILE REFERENCE: Q50979
;  CURRENT APPLICATION NUMBER: US/09/101,272G
;  CURRENT FILING DATE:  1998-07-08
;  PRIOR APPLICATION NUMBER: JP 1059/1996
;  PRIOR FILING DATE: 1996-01-08
;  NUMBER OF SEQ ID NOS: 107
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 95
;   LENGTH: 645
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: ATFHI-CL chimeric protein
;   NAME/KEY: CDS
;   LOCATION: (12)..(614)
;   OTHER INFORMATION:
;   NAME/KEY: mat_peptide
;   LOCATION: (15)..()
;   OTHER INFORMATION:
US-09-101-272G-95

  Query Match             72.3%;  Score 18.8;  DB 4;  Length 645;
  Best Local Similarity   90.9%;  Pred. No. 12;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 CTGGAAGTCTGCCAGTGTCCTT 26
              |||||| ||||||| |||||||
Db        233 CTGGAACTCTGCCACTGTCCTT 254


RESULT 6
US-09-101-272G-97
; Sequence 97, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
;  APPLICANT: Nissin Food Products Co., Ltd.
;  TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
;  FILE REFERENCE: Q50979
;  CURRENT APPLICATION NUMBER: US/09/101,272G
;  CURRENT FILING DATE:  1998-07-08
;  PRIOR APPLICATION NUMBER: JP 1059/1996
;  PRIOR FILING DATE: 1996-01-08
;  NUMBER OF SEQ ID NOS: 107
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 97
;   LENGTH: 666
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: ATFHI-ML chimeric protein
;   NAME/KEY: CDS
;   LOCATION: (12)..(635)
;   OTHER INFORMATION:
;   NAME/KEY: mat_peptide
;   LOCATION: (15)..()
;   OTHER INFORMATION:
US-09-101-272G-97

  Query Match             72.3%;  Score 18.8;  DB 4;  Length 666;
  Best Local Similarity   90.9%;  Pred. No. 12;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 CTGGAAGTCTGCCAGTGTCCTT 26
              |||||| ||||||| |||||||
Db        233 CTGGAACTCTGCCACTGTCCTT 254


RESULT 7
US-08-254-922-1
; Sequence 1, Application US/08254922
; Patent No. 5626841
;  GENERAL INFORMATION:
;    APPLICANT:  Victor Gurewich
;    TITLE OF INVENTION:  USE OF INTRA-PLATELET
;    TITLE OF INVENTION:  UROKINASE-TYPE PLASMINOGEN
;    TITLE OF INVENTION:  ACTIVATORS FOR LONG-TERM
;    TITLE OF INVENTION:  INHIBITION OF THROMBOSIS
;    NUMBER OF SEQUENCES:  1
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      COMPUTER:  IBM PS/2 Model 50Z or 55SX
;      OPERATING SYSTEM:  MS-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/254,922
;      FILING DATE:  June 7, 1994
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/014,207 -
;      FILING DATE:  February 5, 1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  J. Peter Fasse
;      REGISTRATION NUMBER:  32,983
;      REFERENCE/DOCKET NUMBER:  04353/004002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1233
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-254-922-1

  Query Match             72.3%;  Score 18.8;  DB 1;  Length 1233;
  Best Local Similarity   90.9%;  Pred. No. 14;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 CTGGAAGTCTGCCAGTGTCCTT 26
              |||||| ||||||| |||||||
Db        219 CTGGAACTCTGCCACTGTCCTT 240


RESULT 8
US-08-286-748B-1
; Sequence 1, Application US/08286748B
; Patent No. 5759542
```

```
;  GENERAL INFORMATION:
;    APPLICANT:  Victor Gurewich
;    TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE DELIVERY
;    TITLE OF INVENTION:  OF DRUGS BY PLATELETS FOR THE TREATMENT OF
;    TITLE OF INVENTION:  CARDIOVASCULAR AND OTHER DISEASES
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      COMPUTER:  IBM PS/2 Model 50Z or 55SX
;      OPERATING SYSTEM:  MS-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/286,748B
;      FILING DATE:  August 5, 1994
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  J. Peter Fasse
;      REGISTRATION NUMBER:  32,983
;      REFERENCE/DOCKET NUMBER:  04547/013001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1233
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-286-748B-1

  Query Match             72.3%;  Score 18.8;  DB 1;  Length 1233;
  Best Local Similarity   90.9%;  Pred. No. 14;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 CTGGAAGTCTGCCAGTGTCCTT 26
              |||||| ||||||| |||||||
Db        219 CTGGAACTCTGCCACTGTCCTT 240


RESULT 9
US-07-957-039A-7
; Sequence 7, Application US/07957039A
; Patent No. 5389538
;  GENERAL INFORMATION:
;    APPLICANT:  TANABE, TOSHIZUMI
;    APPLICANT:  MORITA, MASANORI
;    APPLICANT:  HIROSE, MASAAKI
;    APPLICANT:  AMASUJI, YASUO
;    TITLE OF INVENTION:  MUTANT HUMAN PROUROKINASE
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sughrue, Mion, Zinn, Macpeak & Seas
;      STREET:  2100 Pennsylvania Avenue
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/957,039A
;      FILING DATE:  06-OCT-1992
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 289257/1991
;      FILING DATE:  07-OCT-1991
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)293-7060
;      TELEFAX:  (202)293-7860
;      TELEX:  6491103
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1236 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  both
;    MOLECULE TYPE:  DNA (genomic)
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  human
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..1233
US-07-957-039A-7

  Query Match             72.3%;  Score 18.8;  DB 1;  Length 1236;
  Best Local Similarity   90.9%;  Pred. No. 14;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 CTGGAAGTCTGCCAGTGTCCTT 26
              |||||| ||||||| |||||||
Db        219 CTGGAACTCTGCCACTGTCCTT 240


RESULT 10
US-08-153-799-17
; Sequence 17, Application US/08153799
; Patent No. 5766883
;  GENERAL INFORMATION:
;    APPLICANT:  Ballance, David J
;    APPLICANT:  Goodey, Andrew R
;    TITLE OF INVENTION:  Polypeptides
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  R Hain Swope, BOC Health Care Inc
;      STREET:  100 Mountain Avenue
;      CITY:  Murray Hill
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07974
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/153,799
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/847975
;      FILING DATE:  06-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 8909916.2
;      FILING DATE:  29-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB90/00650
;      FILING DATE:  26-APR-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/775952
;      FILING DATE:  29-OCT-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Swope, R Hain
```

```
;       REGISTRATION NUMBER:  24864
;       REFERENCE/DOCKET NUMBER:  92H832
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (908) 665 2400
;       TELEFAX:  (908) 771 6159
;       TELEX:  219484
;   INFORMATION FOR SEQ ID NO:  17:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1236 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
;     ORIGINAL SOURCE:
;       ORGANISM:  Homo sapiens
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  1..1236
;       OTHER INFORMATION:  /function= "human mature
;       OTHER INFORMATION:  urokinase-type plasminogen activator (uPA)"
;     FEATURE:
;       NAME/KEY:  primer_bind
;       LOCATION:  13..47
;       OTHER INFORMATION:  /standard_name= "PCR primer binding
;       OTHER INFORMATION:  site"
;     FEATURE:
;       NAME/KEY:  primer_bind
;       LOCATION:  376..418
;       OTHER INFORMATION:  /standard_name= "PCR primer binding
;       OTHER INFORMATION:  site"
US-08-153-799-17

  Query Match             72.3%;  Score 18.8;  DB 1;  Length 1236;
  Best Local Similarity   90.9%;  Pred. No. 14;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 CTGGAAGTCTGCCAGTGTCCTT 26
              |||||| ||||||| |||||||
Db        219 CTGGAACTCTGCCACTGTCCTT 240


RESULT 11
US-09-023-655-927
; Sequence 927, Application US/09023655
; Patent No. 6607879
;  GENERAL INFORMATION:
;    APPLICANT:  Cocks, Benjamin G.
;    APPLICANT:  Susan G. Stuart
;    APPLICANT:  Jeffrey J. Seilhamer
;    TITLE OF INVENTION:  COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
;    TITLE OF INVENTION:  EXPRESSION
;    NUMBER OF SEQUENCES:  1508
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/023,655
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Zeller, Karen J.
;      REGISTRATION NUMBER:  37,071
;      REFERENCE/DOCKET NUMBER:  PA-0001 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 855-0555
;      TELEFAX:  (650) 845-4166
;  INFORMATION FOR SEQ ID NO:  927:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1236 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GENBANK
;      CLONE:  g1311467
US-09-023-655-927

  Query Match             72.3%;  Score 18.8;  DB 4;  Length 1236;
  Best Local Similarity   90.9%;  Pred. No. 14;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 CTGGAAGTCTGCCAGTGTCCTT 26
              |||||| ||||||| |||||||
Db        219 CTGGAACTCTGCCACTGTCCTT 240


RESULT 12
5219569-1
;Patent No. 5219569
;    APPLICANT: BLABER, MICHAEL;HEYNEKER, HERBERT L.;VEHAR,
;GORDON A.
;    TITLE OF INVENTION: PROTEASE RESISTANT UROKINASE
;    NUMBER OF SEQUENCES: 6
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/766,858
;      FILING DATE: 16-AUG-1985
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 725,468
;      FILING DATE: 22-APR-1985
;SEQ ID NO:1:
;      LENGTH: 1372
5219569-1

  Query Match             72.3%;  Score 18.8;  DB 6;  Length 1372;
  Best Local Similarity   90.9%;  Pred. No. 14;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 CTGGAAGTCTGCCAGTGTCCTT 26
              |||||| ||||||| |||||||
Db        355 CTGGAACTCTGCCACTGTCCTT 376


RESULT 13
US-09-643-597-122
; Sequence 122, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Hosken, Nancy
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  McNeill, Patricia D.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.455C11
```

```
;  CURRENT APPLICATION NUMBER: US/09/643,597
;  CURRENT FILING DATE:  2000-08-21
;  NUMBER OF SEQ ID NOS:  369
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 122
;   LENGTH: 1475
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-643-597-122

  Query Match             72.3%;  Score 18.8;  DB 4;  Length 1475;
  Best Local Similarity   90.9%;  Pred. No. 14;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 CTGGAAGTCTGCCAGTGTCCTT 26
              |||||| ||||||| |||||||
Db        359 CTGGAACTCTGCCACTGTCCTT 380


RESULT 14
US-09-480-884A-122
; Sequence 122, Application US/09480884A
; Patent No. 6482597
; GENERAL INFORMATION:
;  APPLICANT: Wang, Tongtong
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Hosken, Nancy A.
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.455C6
;  CURRENT APPLICATION NUMBER: US/09/480,884A
;  CURRENT FILING DATE:  2001-08-27
;  NUMBER OF SEQ ID NOS: 330
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 122
;   LENGTH: 1475
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-480-884A-122

  Query Match             72.3%;  Score 18.8;  DB 4;  Length 1475;
  Best Local Similarity   90.9%;  Pred. No. 14;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 CTGGAAGTCTGCCAGTGTCCTT 26
              |||||| ||||||| |||||||
Db        359 CTGGAACTCTGCCACTGTCCTT 380


RESULT 15
US-09-542-615A-122
; Sequence 122, Application US/09542615A
; Patent No. 6518256
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Hosken, Nancy A.
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION:  COMPOUNDS AND METHODS FOR THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.455C8
;  CURRENT APPLICATION NUMBER: US/09/542,615A
;  CURRENT FILING DATE:  2000-04-14
;  NUMBER OF SEQ ID NOS:  350
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 122
;   LENGTH: 1475
;   TYPE: DNA
;   ORGANISM: Homo sapien
US-09-542-615A-122

  Query Match             72.3%;  Score 18.8;  DB 4;  Length 1475;
  Best Local Similarity   90.9%;  Pred. No. 14;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 CTGGAAGTCTGCCAGTGTCCTT 26
              |||||| ||||||| |||||||
Db        359 CTGGAACTCTGCCACTGTCCTT 380


Search completed: February 20, 2004, 01:45:11
Job time : 6.21902 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 23:23:34 ; Search time 62.4956 Seconds
(without alignments)
1456.787 Million cell updates/sec

Title: US-09-939-853A-141
Perfect score: 26
Sequence: 1 ccttctggaagtctgccagtgtcctt 26

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2308684 seqs, 1750822206 residues

Total number of hits satisfying chosen parameters: 4617368

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	26	100.0	444	9	US-09-867-550-951	Sequence 951, App
c 2	26	100.0	763	9	US-09-867-550-953	Sequence 953, App
c 3	26	100.0	864	10	US-09-814-353-21302	Sequence 21302, A
4	20.4	78.5	611	15	US-10-027-632-195852	Sequence 195852,
5	18.8	72.3	258	15	US-10-233-675A-2	Sequence 2, Appli
6	18.8	72.3	258	15	US-10-233-675A-6	Sequence 6, Appli
7	18.8	72.3	258	15	US-10-233-675A-8	Sequence 8, Appli
8	18.8	72.3	258	15	US-10-233-675A-23	Sequence 23, Appl
9	18.8	72.3	258	15	US-10-233-675A-26	Sequence 26, Appl
10	18.8	72.3	264	9	US-09-880-503-10	Sequence 10, Appl
11	18.8	72.3	288	9	US-09-880-503-18	Sequence 18, Appl
12	18.8	72.3	405	9	US-09-880-503-13	Sequence 13, Appl
13	18.8	72.3	423	9	US-09-984-186-11	Sequence 11, Appl
14	18.8	72.3	423	14	US-10-237-667-11	Sequence 11, Appl
15	18.8	72.3	423	14	US-10-237-708-11	Sequence 11, Appl
16	18.8	72.3	423	14	US-10-237-866-11	Sequence 11, Appl
17	18.8	72.3	423	14	US-10-237-871-11	Sequence 11, Appl
18	18.8	72.3	423	14	US-10-237-624-11	Sequence 11, Appl
19	18.8	72.3	429	9	US-09-880-503-17	Sequence 17, Appl
20	18.8	72.3	482	9	US-09-917-800A-581	Sequence 581, App
21	18.8	72.3	1212	9	US-09-880-503-15	Sequence 15, Appl
22	18.8	72.3	1236	9	US-09-880-503-12	Sequence 12, Appl
23	18.8	72.3	1236	15	US-10-407-821-1	Sequence 1, Appli
24	18.8	72.3	1296	13	US-10-076-421-1	Sequence 1, Appli
25	18.8	72.3	1475	9	US-09-735-705-122	Sequence 122, App
26	18.8	72.3	1475	9	US-09-850-716A-122	Sequence 122, App
27	18.8	72.3	1475	9	US-09-897-778-122	Sequence 122, App
28	18.8	72.3	1475	10	US-09-466-396A-122	Sequence 122, App
29	18.8	72.3	1475	14	US-10-117-982-122	Sequence 122, App
30	18.8	72.3	1475	14	US-10-101-510-159	Sequence 159, App
31	18.8	72.3	1475	15	US-10-313-986-122	Sequence 122, App
32	18.8	72.3	1964	14	US-10-131-985-22	Sequence 22, Appl
33	18.8	72.3	1996	15	US-10-401-077-2	Sequence 2, Appli
34	18.8	72.3	2294	9	US-09-735-705-123	Sequence 123, App
35	18.8	72.3	2294	9	US-09-850-716A-123	Sequence 123, App
36	18.8	72.3	2294	9	US-09-897-778-123	Sequence 123, App
37	18.8	72.3	2294	10	US-09-466-396A-123	Sequence 123, App
38	18.8	72.3	2294	14	US-10-117-982-123	Sequence 123, App
39	18.8	72.3	2294	15	US-10-313-986-123	Sequence 123, App
40	18.8	72.3	2304	14	US-10-171-311-183	Sequence 183, App
41	18.8	72.3	2304	14	US-10-301-822-160	Sequence 160, App
42	18.8	72.3	2304	15	US-10-295-027-413	Sequence 413, App
43	18.8	72.3	2304	15	US-10-295-027-1058	Sequence 1058, Ap
44	18.8	72.3	2304	15	US-10-159-563-121	Sequence 121, App
45	18.8	72.3	2328	14	US-10-247-671-34	Sequence 34, Appl

ALIGNMENTS

RESULT 1
US-09-867-550-951/c
; Sequence 951, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells and F
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 951
; LENGTH: 444
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-867-550-951

Query Match 100.0%; Score 26; DB 9; Length 444;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCTTCTGGAAGTCTGCCAGTGTCCTT 26
||||||||||||||||||||||||||
Db 35 CCTTCTGGAAGTCTGCCAGTGTCCTT 10

RESULT 2
US-09-867-550-953/c
; Sequence 953, Application US/09867550
; Patent No. US20020082206A1

```
; GENERAL INFORMATION:
;  APPLICANT: Leach, Martin D.
;  APPLICANT:  Mehraban, Fuad,
;  APPLICANT:  Conley, Pamela
;  APPLICANT:  Law, Debbie
;  APPLICANT:  Topper, James
;  TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells and
;  TITLE OF INVENTION:  Thereby
;  FILE REFERENCE: 21402-013 (Cura-313)
;  CURRENT APPLICATION NUMBER: US/09/867,550
;  CURRENT FILING DATE:  2001-09-20
;  PRIOR APPLICATION NUMBER: USSN 60/208,427
;  PRIOR FILING DATE: 2000-05-30
;  NUMBER OF SEQ ID NOS: 2125
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 953
;   LENGTH: 763
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-867-550-953

  Query Match            100.0%;  Score 26;  DB 9;  Length 763;
  Best Local Similarity   100.0%;  Pred. No. 0.013;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCTTCTGGAAGTCTGCCAGTGTCCTT 26
              ||||||||||||||||||||||||||
Db        164 CCTTCTGGAAGTCTGCCAGTGTCCTT 139


RESULT 3
US-09-814-353-21302/c
; Sequence 21302, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
;  APPLICANT: Lee, John
;  APPLICANT:  Thompson, Pamela
;  APPLICANT:  Lillie, James
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
;  TITLE OF INVENTION:  IDENTIFICATION, ASSESSMENT, PREVENTION, AND
;  TITLE OF INVENTION:  THERAPY OF OVARIAN CANCER
;  FILE REFERENCE: MRI-006B
;  CURRENT APPLICATION NUMBER: US/09/814,353
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: US 60/191,031
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: US 60/207,124
;  PRIOR FILING DATE: 2000-05-25
;  PRIOR APPLICATION NUMBER: US 60/211,940
;  PRIOR FILING DATE: 2000-06-15
;  PRIOR APPLICATION NUMBER: US 60/216,820
;  PRIOR FILING DATE: 2000-07-07
;  PRIOR APPLICATION NUMBER: US 60/220,661
;  PRIOR FILING DATE: 2000-07-25
;  PRIOR APPLICATION NUMBER: US 60/257,672
;  PRIOR FILING DATE: 2000-12-21
;  NUMBER OF SEQ ID NOS: 22037
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 21302
;   LENGTH: 864
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 1, 2, 3, 32, 862, 863, 864
;   OTHER INFORMATION: n = A,T,C or G
US-09-814-353-21302

  Query Match            100.0%;  Score 26;  DB 10;  Length 864;
  Best Local Similarity   100.0%;  Pred. No. 0.013;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CCTTCTGGAAGTCTGCCAGTGTCCTT 26
              ||||||||||||||||||||||||||
Db        328 CCTTCTGGAAGTCTGCCAGTGTCCTT 303


RESULT 4
US-10-027-632-195852
; Sequence 195852, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 195852
;   LENGTH: 611
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-195852

  Query Match            78.5%;  Score 20.4;  DB 15;  Length 611;
  Best Local Similarity   95.5%;  Pred. No. 5.9;
  Matches   21;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 CTTCTGGAAGTCTGCCAGTGTC 23
              |||||||||||||||||||| |
Db        484 CTTCTGGAAGTCTGCCAGTGCC 505


RESULT 5
US-10-233-675A-2
; Sequence 2, Application US/10233675A
; Publication No. US20030228298A1
; GENERAL INFORMATION:
;  APPLICANT: Nesbit, Mark
;  APPLICANT:  Fong, Timothy
;  APPLICANT:  Brockstedt, Dirk
;  TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
;  TITLE OF INVENTION:  Them To Inhibit Angiogenesis
;  FILE REFERENCE: ST01027
;  CURRENT APPLICATION NUMBER: US/10/233,675A
;  CURRENT FILING DATE:  2002-09-04
;  PRIOR APPLICATION NUMBER: 60/316,300
;  PRIOR FILING DATE: 2001-09-04
;  NUMBER OF SEQ ID NOS: 27
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 258
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: human derived abrogen
US-10-233-675A-2

  Query Match            72.3%;  Score 18.8;  DB 15;  Length 258;
  Best Local Similarity   90.9%;  Pred. No. 33;
```

```
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 CTGGAAGTCTGCCAGTGTCCTT 26
              |||||| ||||||| |||||||
Db         78 CTGGAACTCTGCCACTGTCCTT 99


RESULT 6
US-10-233-675A-6
; Sequence 6, Application US/10233675A
; Publication No. US20030228298A1
; GENERAL INFORMATION:
;  APPLICANT: Nesbit, Mark
;  APPLICANT:  Fong, Timothy
;  APPLICANT:  Brockstedt, Dirk
;  TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
;  TITLE OF INVENTION:  Them To Inhibit Angiogenesis
;  FILE REFERENCE: ST01027
;  CURRENT APPLICATION NUMBER: US/10/233,675A
;  CURRENT FILING DATE:  2002-09-04
;  PRIOR APPLICATION NUMBER: 60/316,300
;  PRIOR FILING DATE: 2001-09-04
;  NUMBER OF SEQ ID NOS: 27
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
;   LENGTH: 258
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: human derived abrogen
US-10-233-675A-6

  Query Match             72.3%;  Score 18.8;  DB 15;  Length 258;
  Best Local Similarity   90.9%;  Pred. No. 33;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 CTGGAAGTCTGCCAGTGTCCTT 26
              |||||| ||||||| |||||||
Db         78 CTGGAACTCTGCCACTGTCCTT 99


RESULT 7
US-10-233-675A-8
; Sequence 8, Application US/10233675A
; Publication No. US20030228298A1
; GENERAL INFORMATION:
;  APPLICANT: Nesbit, Mark
;  APPLICANT:  Fong, Timothy
;  APPLICANT:  Brockstedt, Dirk
;  TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
;  TITLE OF INVENTION:  Them To Inhibit Angiogenesis
;  FILE REFERENCE: ST01027
;  CURRENT APPLICATION NUMBER: US/10/233,675A
;  CURRENT FILING DATE:  2002-09-04
;  PRIOR APPLICATION NUMBER: 60/316,300
;  PRIOR FILING DATE: 2001-09-04
;  NUMBER OF SEQ ID NOS: 27
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
;   LENGTH: 258
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: human derived abrogen
US-10-233-675A-8

  Query Match             72.3%;  Score 18.8;  DB 15;  Length 258;
  Best Local Similarity   90.9%;  Pred. No. 33;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 CTGGAAGTCTGCCAGTGTCCTT 26
              |||||| ||||||| |||||||
Db         78 CTGGAACTCTGCCACTGTCCTT 99


RESULT 8
US-10-233-675A-23
; Sequence 23, Application US/10233675A
; Publication No. US20030228298A1
; GENERAL INFORMATION:
;  APPLICANT: Nesbit, Mark
;  APPLICANT:  Fong, Timothy
;  APPLICANT:  Brockstedt, Dirk
;  TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
;  TITLE OF INVENTION:  Them To Inhibit Angiogenesis
;  FILE REFERENCE: ST01027
;  CURRENT APPLICATION NUMBER: US/10/233,675A
;  CURRENT FILING DATE:  2002-09-04
;  PRIOR APPLICATION NUMBER: 60/316,300
;  PRIOR FILING DATE: 2001-09-04
;  NUMBER OF SEQ ID NOS: 27
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 23
;   LENGTH: 258
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: fragment of human urokinase plasminogen activator
US-10-233-675A-23

  Query Match             72.3%;  Score 18.8;  DB 15;  Length 258;
  Best Local Similarity   90.9%;  Pred. No. 33;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 CTGGAAGTCTGCCAGTGTCCTT 26
              |||||| ||||||| |||||||
Db         78 CTGGAACTCTGCCACTGTCCTT 99


RESULT 9
US-10-233-675A-26
; Sequence 26, Application US/10233675A
; Publication No. US20030228298A1
; GENERAL INFORMATION:
;  APPLICANT: Nesbit, Mark
;  APPLICANT:  Fong, Timothy
;  APPLICANT:  Brockstedt, Dirk
;  TITLE OF INVENTION: Abrogen Polypeptides, Nucleic Acids Encoding Them and Methods for
;  TITLE OF INVENTION:  Them To Inhibit Angiogenesis
;  FILE REFERENCE: ST01027
;  CURRENT APPLICATION NUMBER: US/10/233,675A
;  CURRENT FILING DATE:  2002-09-04
;  PRIOR APPLICATION NUMBER: 60/316,300
;  PRIOR FILING DATE: 2001-09-04
;  NUMBER OF SEQ ID NOS: 27
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26
;   LENGTH: 258
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: fragment of human urokinase plasminogen activator
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (137)..(137)
;   OTHER INFORMATION: n = a or g
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (221)..(221)
;   OTHER INFORMATION: n = a or g
US-10-233-675A-26

  Query Match             72.3%;  Score 18.8;  DB 15;  Length 258;
  Best Local Similarity   90.9%;  Pred. No. 33;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;
```

```
Qy          5 CTGGAAGTCTGCCAGTGTCCTT 26
              |||||| ||||||| |||||||
Db         78 CTGGAACTCTGCCACTGTCCTT 99


RESULT 10
US-09-880-503-10
; Sequence 10, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
;  APPLICANT: CINES, Douglas B
;  APPLICANT:  HIGAZI, Abd Al-Roof
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
;  TITLE OF INVENTION:  TISSUE CONTRACTABILITY
;  FILE REFERENCE: 9596-331
;  CURRENT APPLICATION NUMBER: US/09/880,503
;  CURRENT FILING DATE:  2001-06-13
;  PRIOR APPLICATION NUMBER: US 60/212,847
;  PRIOR FILING DATE: 2000-06-20
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
;   LENGTH: 264
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-880-503-10

  Query Match             72.3%;  Score 18.8;  DB 9;  Length 264;
  Best Local Similarity   90.9%;  Pred. No. 33;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 CTGGAAGTCTGCCAGTGTCCTT 26
              |||||| ||||||| |||||||
Db         78 CTGGAACTCTGCCACTGTCCTT 99


RESULT 11
US-09-880-503-18
; Sequence 18, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
;  APPLICANT: CINES, Douglas B
;  APPLICANT:  HIGAZI, Abd Al-Roof
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
;  TITLE OF INVENTION:  TISSUE CONTRACTABILITY
;  FILE REFERENCE: 9596-331
;  CURRENT APPLICATION NUMBER: US/09/880,503
;  CURRENT FILING DATE:  2001-06-13
;  PRIOR APPLICATION NUMBER: US 60/212,847
;  PRIOR FILING DATE: 2000-06-20
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
;   LENGTH: 288
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-880-503-18

  Query Match             72.3%;  Score 18.8;  DB 9;  Length 288;
  Best Local Similarity   90.9%;  Pred. No. 33;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 CTGGAAGTCTGCCAGTGTCCTT 26
              |||||| ||||||| |||||||
Db         78 CTGGAACTCTGCCACTGTCCTT 99


RESULT 12
US-09-880-503-13
; Sequence 13, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
;  APPLICANT: CINES, Douglas B
;  APPLICANT:  HIGAZI, Abd Al-Roof
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
;  TITLE OF INVENTION:  TISSUE CONTRACTABILITY
;  FILE REFERENCE: 9596-331
;  CURRENT APPLICATION NUMBER: US/09/880,503
;  CURRENT FILING DATE:  2001-06-13
;  PRIOR APPLICATION NUMBER: US 60/212,847
;  PRIOR FILING DATE: 2000-06-20
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
;   LENGTH: 405
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-880-503-13

  Query Match             72.3%;  Score 18.8;  DB 9;  Length 405;
  Best Local Similarity   90.9%;  Pred. No. 34;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 CTGGAAGTCTGCCAGTGTCCTT 26
              |||||| ||||||| |||||||
Db        219 CTGGAACTCTGCCACTGTCCTT 240


RESULT 13
US-09-984-186-11
; Sequence 11, Application US/09984186
; Patent No. US20020151011A1
;  GENERAL INFORMATION:
;       APPLICANT: Fleer, Reinhard
;                  Fournier, Alain
;                  Guitton, Jean-Dominique
;                  Jung, Gerard
;                  Yeh, Patrice
;       TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
;                           PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
;                           CONTAINING SAID POLYPEPTIDES
;       NUMBER OF SEQUENCES: 36
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Rhone-Poulenc Rorer Inc.
;            STREET: 500 Arcola Road, 3C43
;            CITY: Collegeville
;            STATE: PA
;            COUNTRY: USA
;            ZIP: 19426
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: Macintosh
;            OPERATING SYSTEM: System 7.1
;            SOFTWARE: Word 5.1 (PatentIn)
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/984,186
;            FILING DATE: 29-Oct-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US/08/797,689
;            FILING DATE: 31-JAN-1997
;            APPLICATION NUMBER: US 08/256,927
;            FILING DATE: 28-JUL-1994
;            APPLICATION NUMBER: FR 92/01064
;            FILING DATE: 31-JAN-1992
;            APPLICATION NUMBER: PCT/FR93/00085
;            FILING DATE: 28-JAN-1993
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Smith Ph.D., Julie K.
;            REGISTRATION NUMBER: P-38,619
;            REFERENCE/DOCKET NUMBER: ST92006-US
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (610) 454-3839
;            TELEFAX: (610) 454-3808
;  INFORMATION FOR SEQ ID NO: 11:
;       SEQUENCE CHARACTERISTICS:
```

```
;             LENGTH: 423 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;         MOLECULE TYPE: cDNA
;         FEATURE:
;              NAME/KEY:  CDS
;              LOCATION:  3..419
;         SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-09-984-186-11

  Query Match             72.3%;  Score 18.8;  DB 9;  Length 423;
  Best Local Similarity   90.9%;  Pred. No. 34;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 CTGGAAGTCTGCCAGTGTCCTT 26
              |||||| ||||||| |||||||
Db        230 CTGGAACTCTGCCACTGTCCTT 251


RESULT 14
US-10-237-667-11
; Sequence 11, Application US/10237667
; Publication No. US20030022308A1
;   GENERAL INFORMATION:
;         APPLICANT: Fleer, Reinhard
;                    Fournier, Alain
;                    Guitton, Jean-Dominique
;                    Jung, Gerard
;                    Yeh, Patrice
;         TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
;                             PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
;                             CONTAINING SAID POLYPEPTIDES
;         NUMBER OF SEQUENCES: 36
;         CORRESPONDENCE ADDRESS:
;              ADDRESSEE: Rhone-Poulenc Rorer Inc.
;              STREET: 500 Arcola Road, 3C43
;              CITY: Collegeville
;              STATE: PA
;              COUNTRY: USA
;              ZIP: 19426
;         COMPUTER READABLE FORM:
;              MEDIUM TYPE: Floppy disk
;              COMPUTER: Macintosh
;              OPERATING SYSTEM: System 7.1
;              SOFTWARE: Word 5.1 (PatentIn)
;         CURRENT APPLICATION DATA:
;              APPLICATION NUMBER: US/10/237,667
;              FILING DATE: 10-Sep-2002
;              CLASSIFICATION: <Unknown>
;         PRIOR APPLICATION DATA:
;              APPLICATION NUMBER: US/08/797,689
;              FILING DATE: 31-JAN-1997
;              APPLICATION NUMBER: US 08/256,927
;              FILING DATE: 28-JUL-1994
;              APPLICATION NUMBER: FR 92/01064
;              FILING DATE: 31-JAN-1992
;              APPLICATION NUMBER: PCT/FR93/00085
;              FILING DATE: 28-JAN-1993
;         ATTORNEY/AGENT INFORMATION:
;              NAME: Smith Ph.D., Julie K.
;              REGISTRATION NUMBER: P-38,619
;              REFERENCE/DOCKET NUMBER: ST92006-US
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: (610) 454-3839
;              TELEFAX: (610) 454-3808
;   INFORMATION FOR SEQ ID NO: 11:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 423 base pairs
;              TYPE: nucleic acid
;              STRANDEDNESS: double
;              TOPOLOGY: linear
;         MOLECULE TYPE: cDNA
;         FEATURE:
;              NAME/KEY:  CDS
;              LOCATION:  3..419
;         SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-10-237-667-11

  Query Match             72.3%;  Score 18.8;  DB 14;  Length 423;
  Best Local Similarity   90.9%;  Pred. No. 34;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 CTGGAAGTCTGCCAGTGTCCTT 26
              |||||| ||||||| |||||||
Db        230 CTGGAACTCTGCCACTGTCCTT 251


RESULT 15
US-10-237-708-11
; Sequence 11, Application US/10237708
; Publication No. US20030036170A1
;   GENERAL INFORMATION:
;         APPLICANT: Fleer, Reinhard
;                    Fournier, Alain
;                    Guitton, Jean-Dominique
;                    Jung, Gerard
;                    Yeh, Patrice
;         TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
;                             PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
;                             CONTAINING SAID POLYPEPTIDES
;         NUMBER OF SEQUENCES: 36
;         CORRESPONDENCE ADDRESS:
;              ADDRESSEE: Rhone-Poulenc Rorer Inc.
;              STREET: 500 Arcola Road, 3C43
;              CITY: Collegeville
;              STATE: PA
;              COUNTRY: USA
;              ZIP: 19426
;         COMPUTER READABLE FORM:
;              MEDIUM TYPE: Floppy disk
;              COMPUTER: Macintosh
;              OPERATING SYSTEM: System 7.1
;              SOFTWARE: Word 5.1 (PatentIn)
;         CURRENT APPLICATION DATA:
;              APPLICATION NUMBER: US/10/237,708
;              FILING DATE: 10-Sep-2002
;              CLASSIFICATION: <Unknown>
;         PRIOR APPLICATION DATA:
;              APPLICATION NUMBER: US/08/797,689
;              FILING DATE: 31-JAN-1997
;              APPLICATION NUMBER: US 08/256,927
;              FILING DATE: 28-JUL-1994
;              APPLICATION NUMBER: FR 92/01064
;              FILING DATE: 31-JAN-1992
;              APPLICATION NUMBER: PCT/FR93/00085
;              FILING DATE: 28-JAN-1993
;         ATTORNEY/AGENT INFORMATION:
;              NAME: Smith Ph.D., Julie K.
;              REGISTRATION NUMBER: P-38,619
;              REFERENCE/DOCKET NUMBER: ST92006-US
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: (610) 454-3839
;              TELEFAX: (610) 454-3808
;   INFORMATION FOR SEQ ID NO: 11:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 423 base pairs
;              TYPE: nucleic acid
;              STRANDEDNESS: double
;              TOPOLOGY: linear
;         MOLECULE TYPE: cDNA
;         FEATURE:
;              NAME/KEY:  CDS
;              LOCATION:  3..419
;         SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-10-237-708-11
```

Query Match 72.3%; Score 18.8; DB 14; Length 423;
Best Local Similarity 90.9%; Pred. No. 34;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          5 CTGGAAGTCTGCCAGTGTCCTT 26
              |||||| ||||||| |||||||
Db        230 CTGGAACTCTGCCACTGTCCTT 251
```

Search completed: February 20, 2004, 05:59:21
Job time : 64.4956 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 21:21:18 ; Search time 96.3181 Seconds
(without alignments)
9899.970 Million cell updates/sec

Title: US-09-939-853A-142
Perfect score: 22
Sequence: 1 tgagagagttctgggtgtccta 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	22	100.0	22	6	AX443201	AX443201 Sequence
	2	22	100.0	1183	6	AX443133	AX443133 Sequence
c	3	22	100.0	1183	6	AX443135	AX443135 Sequence
	4	22	100.0	2538	9	BC042041	BC042041 Homo sapi
	5	22	100.0	2567	6	AX452880	AX452880 Sequence
	6	22	100.0	2788	6	AX780857	AX780857 Sequence
c	7	22	100.0	66741	9	HS460J8	AL031662 Human DNA
	8	22	100.0	145833	2	AC026539	AC026539 Homo sapi
	9	19.4	88.2	110000	9	AC005079_2	Continuation (3 of
c	10	19.4	88.2	170128	2	AC005059	AC005059 Homo sapi
c	11	19.4	88.2	224064	2	AC005055	AC005055 Homo sapi
	12	18.8	85.5	3521	3	AY058623	AY058623 Drosophil
c	13	18.8	85.5	26121	2	AC015346	AC015346 Drosophil
c	14	18.8	85.5	177917	9	AC025824	AC025824 Homo sapi
c	15	18.8	85.5	180213	3	AC010847	AC010847 Drosophil
c	16	18.8	85.5	182623	3	AC011071	AC011071 Drosophil
c	17	18.8	85.5	186540	9	AC012478	AC012478 Homo sapi
	18	18.8	85.5	186674	9	AC091544	AC091544 Homo sapi
	19	18.8	85.5	188365	2	AC116758	AC116758 Mus muscu
c	20	18.8	85.5	346474	3	AE003512	AE003512 Drosophil
	21	18.4	83.6	76640	10	AL928947	AL928947 Mouse DNA
	22	18.4	83.6	178980	2	AC120340	AC120340 Mus muscu
	23	18.4	83.6	181791	2	AC101938	AC101938 Mus muscu
	24	18.4	83.6	193252	2	AC122504	AC122504 Mus muscu
	25	18.4	83.6	245603	2	AC124616	AC124616 Mus muscu
	26	18	81.8	226598	2	AC099631	AC099631 Mus muscu
c	27	18	81.8	271101	2	AC132376	AC132376 Mus muscu
c	28	17.8	80.9	1177	3	AK114842	AK114842 Ciona int
c	29	17.8	80.9	2064	6	AX834238	AX834238 Sequence
c	30	17.8	80.9	2064	9	AK096566	AK096566 Homo sapi
c	31	17.8	80.9	45126	9	AL590402	AL590402 Human DNA
c	32	17.8	80.9	70328	9	AL513472	AL513472 Human DNA
	33	17.8	80.9	91125	9	AC004899	AC004899 Homo sapi
	34	17.8	80.9	148869	9	AC024093	AC024093 Homo sapi
	35	17.8	80.9	155085	9	AC010412	AC010412 Homo sapi
c	36	17.8	80.9	162190	2	AC020592	AC020592 Homo sapi
	37	17.8	80.9	168670	2	AC139753	AC139753 Mus muscu
c	38	17.8	80.9	169072	5	AL929396	AL929396 Zebrafish
c	39	17.8	80.9	171343	2	AC016676	AC016676 Homo sapi
c	40	17.8	80.9	172980	10	AC126671	AC126671 Mus muscu
c	41	17.8	80.9	197012	10	AL669823	AL669823 Mouse DNA
	42	17.8	80.9	202681	2	AC111090	AC111090 Mus muscu
	43	17.8	80.9	204131	2	AC128898	AC128898 Rattus no
	44	17.8	80.9	210194	10	AL805896	AL805896 Mouse DNA
	45	17.8	80.9	216589	9	AY129950	AY129950 Homo sapi

ALIGNMENTS

RESULT 1
AX443201
LOCUS AX443201 22 bp DNA linear PAT 02-JUL-2002
DEFINITION Sequence 142 from Patent WO0216599.
ACCESSION AX443201
VERSION AX443201.1 GI:21690596
KEYWORDS .
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1
AUTHORS Burgess,C.E., Conley,P.B., Grosse,W.M., Hart,M., Kekuda,R., Shimkets,R.A., Spytek,K.A., Szekeres,E.S., Tomlinson,J.E., Topper,J.N. and Yang,R.B.
TITLE Proteins and nucleic acids encoding same

```
  JOURNAL   Patent: WO 0216599-A 142 28-FEB-2002;
            Curagen Corporation (US) ; COR THERAPEUTICS, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="oligonucleotide primer"
ORIGIN

  Query Match           100.0%;  Score 22;  DB 6;  Length 22;
  Best Local Similarity  100.0%;  Pred. No. 0.27;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTCCTA 22
              ||||||||||||||||||||||
Db          1 TGAGAGAGTTCTGGGTGTCCTA 22


RESULT 2
AX443133
LOCUS       AX443133               1183 bp    DNA     linear   PAT 02-JUL-2002
DEFINITION  Sequence 74 from Patent WO0216599.
ACCESSION   AX443133
VERSION     AX443133.1  GI:21690555
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Burgess,C.E., Conley,P.B., Grosse,W.M., Hart,M., Kekuda,R.,
            Shimkets,R.A., Spytek,K.A., Szekeres,E.S., Tomlinson,J.E.,
            Topper,J.N. and Yang,R.B.
  TITLE     Proteins and nucleic acids encoding same
  JOURNAL   Patent: WO 0216599-A 74 28-FEB-2002;
            Curagen Corporation (US) ; COR THERAPEUTICS, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .1183
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           100.0%;  Score 22;  DB 6;  Length 1183;
  Best Local Similarity  100.0%;  Pred. No. 0.36;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTCCTA 22
              ||||||||||||||||||||||
Db        224 TGAGAGAGTTCTGGGTGTCCTA 245


RESULT 3
AX443135/c
LOCUS       AX443135               1183 bp    DNA     linear   PAT 02-JUL-2002
DEFINITION  Sequence 76 from Patent WO0216599.
ACCESSION   AX443135
VERSION     AX443135.1  GI:21690556
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Burgess,C.E., Conley,P.B., Grosse,W.M., Hart,M., Kekuda,R.,
            Shimkets,R.A., Spytek,K.A., Szekeres,E.S., Tomlinson,J.E.,
            Topper,J.N. and Yang,R.B.
  TITLE     Proteins and nucleic acids encoding same
  JOURNAL   Patent: WO 0216599-A 76 28-FEB-2002;
            Curagen Corporation (US) ; COR THERAPEUTICS, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .1183
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match           100.0%;  Score 22;  DB 6;  Length 1183;
  Best Local Similarity  100.0%;  Pred. No. 0.36;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTCCTA 22
              ||||||||||||||||||||||
Db        960 TGAGAGAGTTCTGGGTGTCCTA 939


RESULT 4
BC042041
LOCUS       BC042041               2538 bp    mRNA    linear   PRI 07-OCT-2003
DEFINITION  Homo sapiens Src-like-adaptor 2, transcript variant 1, mRNA (cDNA
            clone MGC:49845 IMAGE:4429896), complete cds.
ACCESSION   BC042041
VERSION     BC042041.1  GI:27469842
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2538)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2  (bases 1 to 2538)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-DEC-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 88 Row: a Column: 20
```

```
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 28416422.
FEATURES             Location/Qualifiers
     source          1. .2538
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:49845 IMAGE:4429896"
                     /tissue_type="Prostate, adenocarcinoma."
                     /clone_lib="NIH_MGC_91"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene            1. .2538
                     /gene="SLA2"
                     /note="synonyms: FLJ21992, SLAP-2, C20orf156, MGC49845,
                     SLAP2"
                     /db_xref="LocusID:84174"
                     /db_xref="MIM:606577"
     CDS             363. .1148
                     /codon_start=1
                     /product="Src-like-adaptor 2, isoform a"
                     /protein_id="AAH42041.1"
                     /db_xref="GI:27469843"
                     /db_xref="LocusID:84174"
                     /translation="MGSLPSRRKSLPSPSLSSSVQGQGPVTMEAERSKATAVALGSFP
                     AGGPAELSLRLGEPLTIVSEDGDWWTVLSEVSGREYNIPSVHVAKVSHGWLYEGLSRE
                     KAEELLLLPGNPGGAFLIRESQTRRGSYSLSVRLSRPASWDRIRHYRIHCLDNGWLYI
                     SPRLTFPSLQALVDHYSELADDICCLLKEPCVLQRAGPLPGKDIPLPVTVQRTPLNWK
                     ELDSSLLFSEAATGEESLLSEGLRESLSFYISLNDEAVSLDDA"
     misc_feature    468. .632
                     /note="SH3; Region: SH3 domain. SH3 (Src homology 3)
                     domains are often indicative of a protein involved in
                     signal transduction related to cytoskeletal organization.
                     First described in the Src cytoplasmic tyrosine kinase.
                     The structure is a partly opened beta barrel"
                     /db_xref="CDD:pfam00018"
     misc_feature    642. .890
                     /note="SH2; Region: SH2 domain"
                     /db_xref="CDD:pfam00017"
ORIGIN

  Query Match             100.0%;  Score 22;  DB 9;  Length 2538;
  Best Local Similarity   100.0%;  Pred. No. 0.38;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTCCTA 22
              ||||||||||||||||||||||
Db        189 TGAGAGAGTTCTGGGTGTCCTA 210


RESULT 5
AX452880
LOCUS       AX452880               2567 bp    DNA     linear   PAT 06-JUL-2002
DEFINITION  Sequence 1 from Patent WO0242457.
ACCESSION   AX452880
VERSION     AX452880.1  GI:21712520
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Chang,H., Yang,W.P., Wu,Y., Whitney,G.S., Perez-Villar,J.J. and
            Kanner,S.B.
  TITLE     Cloning and expression of human slap-2: a novel sh2/sh3
            domain-containing human slap homologue having immune cell-specific
            expression
  JOURNAL   Patent: WO 0242457-A 1 30-MAY-2002;
            Bristol-Myers Squibb Co. (US)
FEATURES             Location/Qualifiers
     source          1. .2567
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             100.0%;  Score 22;  DB 6;  Length 2567;
  Best Local Similarity   100.0%;  Pred. No. 0.38;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTCCTA 22
              ||||||||||||||||||||||
Db        241 TGAGAGAGTTCTGGGTGTCCTA 262


RESULT 6
AX780857
LOCUS       AX780857               2788 bp    DNA     linear   PAT 14-JUL-2003
DEFINITION  Sequence 3014 from Patent WO03039443.
ACCESSION   AX780857
VERSION     AX780857.1  GI:32697851
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Haferlach,T., Schoch,C., Kern,W., Kohlmann,A., Schnittger,S.,
            Dugas,M., Eils,R., Brors,B. and Mergenthaler,S.
  TITLE     Novel genetic markers for leukemias
  JOURNAL   Patent: WO 03039443-A 3014 15-MAY-2003;
            Deutsches Krebsforschungszentrum (DE) ;
            Ludwig-Maximilian-Universitaet Muenchen (DE) ; Haferlach, Torsten,
            PD Dr. Dr. (DE) ; Schoch, Claudia (DE) ; Kern, Wolfgang (DE)
FEATURES             Location/Qualifiers
     source          1. .2788
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             100.0%;  Score 22;  DB 6;  Length 2788;
  Best Local Similarity   100.0%;  Pred. No. 0.38;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTCCTA 22
              ||||||||||||||||||||||
Db        213 TGAGAGAGTTCTGGGTGTCCTA 234


RESULT 7
HS460J8/c
LOCUS       HS460J8                66741 bp    DNA     linear   PRI 23-JUL-2001
DEFINITION  Human DNA sequence from clone RP3-460J8 on chromosome
            20q11.21-11.23 Contains the 3' end of the gene for a novel protein
            similar to N-myc downstream regulated (NDRG1) the 5' end of a gene
            encoding a novel protein tyrosine kinase, ESTs, STSs and GSSs,
            complete sequence.
ACCESSION   AL031662
VERSION     AL031662.26  GI:9716901
KEYWORDS    HTG; NDRG1; SH2 domain.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 66741)
  AUTHORS   Skuce,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Aug 7, 2000 this sequence version replaced gi:6425549.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
```

corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/Chr20
IMPORTANT: This sequence is not the entire insert of clone RP3-460J8 It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.
The true right end of clone RP3-460J8 is at 66741 in this sequence.
The true left end of clone RP3-469A13 is at 41767 in this sequence.
The true right end of clone RP5-977B1 is at 100 in this sequence.
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP3-460J8 is from the library RPCI-3 constructed by the group of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pCYPAC2.

FEATURES	Location/Qualifiers
source	1. .66741
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="20"
	/map="q11.21-11.23"
	/clone="RP3-460J8"
	/clone_lib="RPCI-3"
misc_feature	complement(50. .544)
	/note="match: GSS: Em:B45150"
gene	complement(240. .7128)
	/gene="dJ977B1.1"
mRNA	complement(join(<240. .339,6995. .7128))
	/gene="dJ977B1.1"
	/product="dJ460J8.2 (novel protein tyrosine kinase with Src homology 2 (SH2) domain)"
	/note="match: cDNAs: Em:AK025645 match: ESTs: Em:BG178487"
	/evidence=not_experimental
CDS	complement(join(<240. .339,6995. .7085))
	/gene="dJ977B1.1"
	/note="Continues in Em:AL050318 as dJ977B1.2"
	/codon_start=1
	/evidence=not_experimental
	/product="dJ460J8.2 (novel protein tyrosine kinase with Src homology 2 (SH2) domain)"
	/protein_id="CAC44645.1"
	/db_xref="GI:15020830"
	/db_xref="GOA:Q9H6Q3"
	/db_xref="SWISS-PROT:Q9H6Q3"
	/translation="MGSLPSRRKSLPSPSLSSSVQGQGPVTMEAERSKATAVALGSFP AGGPAELSLRLGEPLTIVS"
repeat_region	536. .565
	/note="15 copies 2 mer ag 93% conserved"
repeat_region	2375. .2402
	/note="14 copies 2 mer ta 100% conserved"
misc_feature	2783. .3155
	/note="match: GSS: Em:AQ807191"
misc_feature	complement(6983. .7482)
	/note="match: GSS: Em:AQ556467"
misc_feature	complement(6985. .7492)
	/note="match: GSS: Em:AQ556478"
misc_feature	7486. .7925
	/note="match: STS: Em:HS427J1S"
repeat_region	10615. .10640
	/note="13 copies 2 mer tt 92% conserved"
repeat_region	13593. .13961
	/note="MER21B repeat: matches 422. .785 of consensus"
repeat_region	14248. .14297
	/note="MER21B repeat: matches 374. .422 of consensus"
repeat_region	14594. .14614
	/note="MER21B repeat: matches 355. .374 of consensus"
repeat_region	14746. .15078
	/note="MER21B repeat: matches 1. .355 of consensus"
polyA_site	complement(17516)
misc_feature	17519. .17686
	/note="match: GSS: Em:AZ067993"
polyA_site	complement(17521)
polyA_site	complement(17522)
gene	complement(17523. .54534)
	/gene="dJ469A13.3"
mRNA	complement(join(<17523. .19451,20551. .20589,22110. .22161, 22247. .22282,26087. .26134,30796. .30847,30986. .31051, 32039. .32142,37091. .37147,46566. .46652,48268. .48328, 50163. .50225,53242. .53362,54429. .>54534))
	/gene="dJ469A13.3"
	/product="dJ460J8.1 (continued from dJ469A13.3 in Em:AL132768)"
	/note="match: cDNAs: Em:AB033922 Em:U52073 Em:D87953 Em:X92845 Em:AF004162 Em:U60593 Em:AF045564 Em:AB033921 Em:AF159092 Em:AF147402 Em:M59814 match: ESTs: Em:AA718726 Em:AA039000 Em:AW003952 Em:W89263 Em:AA325826 Em:AI230982 Em:AA162360 Em:AA445016 Em:AV002395 Em:T85147 Em:AI786615 Em:AA113437 Em:AI004026 Em:T88705 Em:AI786673 Em:AI181197 Em:AA764653 Em:AA316771 Em:AU035165 Em:AV002368"
	/evidence=not_experimental
polyA_signal	complement(17542. .17547)
	/gene="dJ469A13.3"
misc_feature	17555. .17811
	/note="match: STS: Em:G19945"
misc_feature	join(19202. .19451,22110. .22282,26087. .26127,30796. .30847)
	/note="match: STS: Em:G23762"
CDS	complement(join(19270. .19451,20551. .20589,22110. .22161, 22247. .22282,26087. .26134,30796. .30847,30986. .31051, 32039. .32142,37091. .37147,46566. .46652,48268. .48328, 50163. .50225,53242. .53362,54429. .>54534))
	/gene="dJ469A13.3"
	/note="novel protein (FLJ13556) similar to N-myc downstream regulated (NDRG1) match: proteins: Sw:Q62433 Sw:Q92597 Tr:Q9Z2L9 Sw:P97862"
	/codon_start=1
	/evidence=not_experimental
	/product="dJ460J8.1 (continued from dJ469A13.3 in Em:AL132768)"
	/protein_id="CAB65625.1"
	/db_xref="GI:6687781"
	/db_xref="GOA:Q9UGV2"
	/db_xref="SWISS-PROT:Q9UGV2"
	/translation="EHDIETTHGVVHVTIRGLPKGNRPVILTYHDIGLNHKSCFNAFF NFEDMQEITQHFAVCHVDAPGQQEGAPSFPTGYQYPTMDELAEMLPPVLTHLSLKSII GIGVGAGAYILSRFALNHPELVEGLVLINVDPCAKGWIDWAASKLSGLTTNVVDIILA HHFGQEELQANLDLIQTYRMHIAQDINQDNLQLFLNSYNGRRDLEIERPILGQNDNKS KTLKCSTLLVVGDNSPAVEAVVECNSRLNPINTTLLKMADCGGLPQVVQPGKLTEAFK YFLQGMGYIPYVQLSHLSTESVPSASMTRLARSRTHSTSSSLGSGESPFSRSVTSNQS DGTQESCESPDVLDRHQTMEVSC"
misc_feature	complement(24246. .24780)
	/gene="dJ469A13.3"
	/note="match: GSS: Em:AQ592789"
misc_feature	complement(27980. .28556)
	/gene="dJ469A13.3"
	/note="match: GSS: Em:AQ308867"
repeat_region	34313. .34420
	/note="HY1 repeat: matches 1. .109 of consensus"
repeat_region	34505. .34560
	/note="28 copies 2 mer ta 78% conserved"
repeat_region	34564. .34611

```
                     /note="24 copies 2 mer ta 79% conserved"
     repeat_region   42553. .43658
                     /note="MER52C repeat: matches 1. .1278 of consensus"
     misc_feature    47989. .48470
                     /note="match: GSS: Em:AQ672417"
     repeat_region   49391. .49480
                     /note="45 copies 2 mer ta 76% conserved"
     misc_feature    52922. .53120
                     /note="match: STS: Em:G04621"
     misc_feature    complement(53426. .54032)
                     /gene="dJ469A13.3"
                     /note="match: GSS: Em:AQ314824
                     match: STS: Em:G55853"
     misc_feature    54054. .54477
                     /note="match: GSS: Em:AQ310681"
     misc_feature    57509. .57838
                     /note="match: STS: Em:G07504"
     misc_feature    complement(57696. .58013)
                     /note="match: STS: Em:G07632 Em:G07634"
     repeat_region   58856. .59001
                     /note="73 copies 2 mer at 83% conserved"
     misc_feature    complement(58930. .59281)
                     /note="match: GSS: Em:AQ067563"
     repeat_region   59062. .59236
                     /note="TIGGER2 repeat: matches 2541. .2714 of consensus"
     repeat_region   59738. .60495
                     /note="TIGGER2 repeat: matches 1780. .2541 of consensus"
     repeat_region   60719. .61831
                     /note="MER11C repeat: matches 1. .1071 of consensus"
     repeat_region   62547. .63174

  Query Match             100.0%;  Score 22;  DB 9;  Length 66741;
  Best Local Similarity   100.0%;  Pred. No. 0.47;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTCCTA 22
              ||||||||||||||||||||||
Db      11717 TGAGAGAGTTCTGGGTGTCCTA 11696


RESULT 8
AC026539
LOCUS       AC026539              145833 bp    DNA     linear   HTG 27-APR-2000
DEFINITION  Homo sapiens chromosome 20 clone RP11-712N14 map 20, WORKING DRAFT
            SEQUENCE, 37 unordered pieces.
ACCESSION   AC026539
VERSION     AC026539.2  GI:7656813
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 145833)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 20, clone RP11-712N14
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 145833)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-MAR-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Apr 27, 2000 this sequence version replaced gi:7283243.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L7115
                Center clone name: 712_N_14
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 125577 bases at least Q40
                Consensus quality: 135703 bases at least Q30
                Consensus quality: 139593 bases at least Q20
                Insert size: 182000; agarose-fp
                Insert size: 142233; sum-of-contigs
                Quality coverage: 2.6 in Q20 bases; agarose-fp
                Quality coverage: 3.3 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 37 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1      1214: contig of 1214 bp in length
            *     1215      1314: gap of 100 bp
            *     1315      2673: contig of 1359 bp in length
            *     2674      2773: gap of 100 bp
            *     2774      4520: contig of 1747 bp in length
            *     4521      4620: gap of 100 bp
            *     4621      5961: contig of 1341 bp in length
            *     5962      6061: gap of 100 bp
            *     6062      7719: contig of 1658 bp in length
            *     7720      7819: gap of 100 bp
            *     7820      9799: contig of 1980 bp in length
            *     9800      9899: gap of 100 bp
            *     9900     11434: contig of 1535 bp in length
            *    11435     11534: gap of 100 bp
            *    11535     14382: contig of 2848 bp in length
            *    14383     14482: gap of 100 bp
            *    14483     17148: contig of 2666 bp in length
            *    17149     17248: gap of 100 bp
            *    17249     19239: contig of 1991 bp in length
            *    19240     19339: gap of 100 bp
            *    19340     21102: contig of 1763 bp in length
            *    21103     21202: gap of 100 bp
            *    21203     23371: contig of 2169 bp in length
            *    23372     23471: gap of 100 bp
            *    23472     25782: contig of 2311 bp in length
            *    25783     25882: gap of 100 bp
            *    25883     28824: contig of 2942 bp in length
            *    28825     28924: gap of 100 bp
            *    28925     31619: contig of 2695 bp in length
            *    31620     31719: gap of 100 bp
            *    31720     34679: contig of 2960 bp in length
            *    34680     34779: gap of 100 bp
            *    34780     37527: contig of 2748 bp in length
```

```
            *    37528    37627: gap of 100 bp
            *    37628    40146: contig of 2519 bp in length
            *    40147    40246: gap of 100 bp
            *    40247    43743: contig of 3497 bp in length
            *    43744    43843: gap of 100 bp
            *    43844    47133: contig of 3290 bp in length
            *    47134    47233: gap of 100 bp
            *    47234    51023: contig of 3790 bp in length
            *    51024    51123: gap of 100 bp
            *    51124    54935: contig of 3812 bp in length
            *    54936    55035: gap of 100 bp
            *    55036    59553: contig of 4518 bp in length
            *    59554    59653: gap of 100 bp
            *    59654    64428: contig of 4775 bp in length
            *    64429    64528: gap of 100 bp
            *    64529    69211: contig of 4683 bp in length
            *    69212    69311: gap of 100 bp
            *    69312    72901: contig of 3590 bp in length
            *    72902    73001: gap of 100 bp
            *    73002    76724: contig of 3723 bp in length
            *    76725    76824: gap of 100 bp
            *    76825    81179: contig of 4355 bp in length
            *    81180    81279: gap of 100 bp
            *    81280    87009: contig of 5730 bp in length
            *    87010    87109: gap of 100 bp
            *    87110    90855: contig of 3746 bp in length
            *    90856    90955: gap of 100 bp
            *    90956    96520: contig of 5565 bp in length
            *    96521    96620: gap of 100 bp
            *    96621   102321: contig of 5701 bp in length
            *   102322   102421: gap of 100 bp
            *   102422   108293: contig of 5872 bp in length
            *   108294   108393: gap of 100 bp
            *   108394   116689: contig of 8296 bp in length
            *   116690   116789: gap of 100 bp
            *   116790   125264: contig of 8475 bp in length
            *   125265   125364: gap of 100 bp
            *   125365   136354: contig of 10990 bp in length
            *   136355   136454: gap of 100 bp
            *   136455   145833: contig of 9379 bp in length.
FEATURES             Location/Qualifiers
     source          1. .145833
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /map="20"
                     /clone="RP11-712N14"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .1214
                     /note="assembly_fragment"
     misc_feature    1315. .2673
                     /note="assembly_fragment"
     misc_feature    2774. .4520
                     /note="assembly_fragment"
     misc_feature    4621. .5961
                     /note="assembly_fragment"
     misc_feature    6062. .7719
                     /note="assembly_fragment"
     misc_feature    7820. .9799
                     /note="assembly_fragment"
     misc_feature    9900. .11434
                     /note="assembly_fragment"
     misc_feature    11535. .14382
                     /note="assembly_fragment"
     misc_feature    14483. .17148
                     /note="assembly_fragment"
     misc_feature    17249. .19239
                     /note="assembly_fragment"
     misc_feature    19340. .21102
                     /note="assembly_fragment"
     misc_feature    21203. .23371
                     /note="assembly_fragment"
     misc_feature    23472. .25782
                     /note="assembly_fragment"
     misc_feature    25883. .28824
                     /note="assembly_fragment"
     misc_feature    28925. .31619
                     /note="assembly_fragment"
     misc_feature    31720. .34679
                     /note="assembly_fragment"
     misc_feature    34780. .37527
                     /note="assembly_fragment"
     misc_feature    37628. .40146
                     /note="assembly_fragment"
     misc_feature    40247. .43743
                     /note="assembly_fragment"
     misc_feature    43844. .47133
                     /note="assembly_fragment"
     misc_feature    47234. .51023
                     /note="assembly_fragment"
     misc_feature    51124. .54935
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
     misc_feature    55036. .59553

  Query Match             100.0%;  Score 22;  DB 2;  Length 145833;
  Best Local Similarity   100.0%;  Pred. No. 0.5;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTCCTA 22
              ||||||||||||||||||||||
Db      99968 TGAGAGAGTTCTGGGTGTCCTA 99989


RESULT 9
AC005079_2
WPCOMMENT
  Sequence split into 5 fragments  LOCUS AC005079 Accession AC005079
     Fragment Name      Begin     End
     AC005079_0             1    110000
     AC005079_1        100001    210000
     AC005079_2        200001    310000
     AC005079_3        300001    410000
     AC005079_4        400001    415268
Continuation (3 of 5) of AC005079 from base 200001 (AC005079 Homo sapiens BAC clone CTA-

  Query Match             88.2%;  Score 19.4;  DB 9;  Length 110000;
  Best Local Similarity   95.2%;  Pred. No. 15;
  Matches   20;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTCCT 21
              |||||||||||||||||| |||
Db      89559 TGAGAGAGTTCTGGGTGGCCT 89579


RESULT 10
AC005059/c
LOCUS       AC005059              170128 bp    DNA     linear   HTG 13-MAR-1999
DEFINITION  Homo sapiens clone RG074A24, *** SEQUENCING IN PROGRESS ***, 25
            unordered pieces.
ACCESSION   AC005059
VERSION     AC005059.1  GI:3212932
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 170128)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 170128)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
```

```
  JOURNAL   Submitted (12-JUN-1998) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     * NOTE: This is a 'working draft' sequence. It currently
            * consists of 25 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1662: contig of 1662 bp in length
            *     1663     1679: gap of unknown length
            *     1680     3248: contig of 1569 bp in length
            *     3249     3265: gap of unknown length
            *     3266     4902: contig of 1637 bp in length
            *     4903     4919: gap of unknown length
            *     4920     6551: contig of 1632 bp in length
            *     6552     6568: gap of unknown length
            *     6569     8090: contig of 1522 bp in length
            *     8091     8107: gap of unknown length
            *     8108     9832: contig of 1725 bp in length
            *     9833     9849: gap of unknown length
            *     9850    11514: contig of 1665 bp in length
            *    11515    11531: gap of unknown length
            *    11532    13880: contig of 2349 bp in length
            *    13881    13897: gap of unknown length
            *    13898    16342: contig of 2445 bp in length
            *    16343    16359: gap of unknown length
            *    16360    20688: contig of 4329 bp in length
            *    20689    20705: gap of unknown length
            *    20706    24570: contig of 3865 bp in length
            *    24571    24587: gap of unknown length
            *    24588    28578: contig of 3991 bp in length
            *    28579    28595: gap of unknown length
            *    28596    35314: contig of 6719 bp in length
            *    35315    35331: gap of unknown length
            *    35332    39541: contig of 4210 bp in length
            *    39542    39558: gap of unknown length
            *    39559    46401: contig of 6843 bp in length
            *    46402    46418: gap of unknown length
            *    46419    53312: contig of 6894 bp in length
            *    53313    53329: gap of unknown length
            *    53330    60264: contig of 6935 bp in length
            *    60265    60281: gap of unknown length
            *    60282    66853: contig of 6572 bp in length
            *    66854    66870: gap of unknown length
            *    66871    75039: contig of 8169 bp in length
            *    75040    75056: gap of unknown length
            *    75057    84782: contig of 9726 bp in length
            *    84783    84799: gap of unknown length
            *    84800    97374: contig of 12575 bp in length
            *    97375    97391: gap of unknown length
            *    97392   109799: contig of 12408 bp in length
            *   109800   109816: gap of unknown length
            *   109817   124892: contig of 15076 bp in length
            *   124893   124909: gap of unknown length
            *   124910   143882: contig of 18973 bp in length
            *   143883   143899: gap of unknown length
            *   143900   170128: contig of 26229 bp in length.
FEATURES             Location/Qualifiers
     source          1. .170128
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="RG074A24"
ORIGIN

  Query Match           88.2%;  Score 19.4;  DB 2;  Length 170128;
  Best Local Similarity   95.2%;  Pred. No. 16;
  Matches   20;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTCCT 21
                  ||||||||||||||||||| |||
Db      48353 TGAGAGAGTTCTGGGTGGCCT 48333


RESULT 11
AC005055/c
LOCUS       AC005055               224064 bp    DNA     linear   HTG 13-MAR-1999
DEFINITION  Homo sapiens clone RG050N15, *** SEQUENCING IN PROGRESS ***, 26
            unordered pieces.
ACCESSION   AC005055
VERSION     AC005055.1  GI:3212936
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 224064)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 224064)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JUN-1998) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     * NOTE: This is a 'working draft' sequence. It currently
            * consists of 26 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1567: contig of 1567 bp in length
            *     1568     1584: gap of unknown length
            *     1585     3757: contig of 2173 bp in length
            *     3758     3774: gap of unknown length
            *     3775     5645: contig of 1871 bp in length
            *     5646     5662: gap of unknown length
            *     5663     8218: contig of 2556 bp in length
            *     8219     8235: gap of unknown length
            *     8236    10568: contig of 2333 bp in length
            *    10569    10585: gap of unknown length
            *    10586    12673: contig of 2088 bp in length
            *    12674    12690: gap of unknown length
            *    12691    14458: contig of 1768 bp in length
            *    14459    14475: gap of unknown length
            *    14476    17399: contig of 2924 bp in length
            *    17400    17416: gap of unknown length
            *    17417    20845: contig of 3429 bp in length
            *    20846    20862: gap of unknown length
            *    20863    24474: contig of 3612 bp in length
            *    24475    24491: gap of unknown length
            *    24492    30780: contig of 6289 bp in length
            *    30781    30797: gap of unknown length
            *    30798    36366: contig of 5569 bp in length
            *    36367    36383: gap of unknown length
            *    36384    41721: contig of 5338 bp in length
            *    41722    41738: gap of unknown length
            *    41739    49140: contig of 7402 bp in length
            *    49141    49157: gap of unknown length
            *    49158    54878: contig of 5721 bp in length
            *    54879    54895: gap of unknown length
            *    54896    60864: contig of 5969 bp in length
            *    60865    60881: gap of unknown length
            *    60882    72111: contig of 11230 bp in length
            *    72112    72128: gap of unknown length
            *    72129    81775: contig of 9647 bp in length
            *    81776    81792: gap of unknown length
            *    81793    92380: contig of 10588 bp in length
            *    92381    92397: gap of unknown length
            *    92398   106304: contig of 13907 bp in length
```

```
            *   106305   106321: gap of unknown length
            *   106322   119608: contig of 13287 bp in length
            *   119609   119625: gap of unknown length
            *   119626   136119: contig of 16494 bp in length
            *   136120   136136: gap of unknown length
            *   136137   157328: contig of 21192 bp in length
            *   157329   157345: gap of unknown length
            *   157346   176995: contig of 19650 bp in length
            *   176996   177012: gap of unknown length
            *   177013   198965: contig of 21953 bp in length
            *   198966   198982: gap of unknown length
            *   198983   224064: contig of 25082 bp in length.
FEATURES             Location/Qualifiers
     source          1. .224064
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="RG050N15"
ORIGIN

  Query Match             88.2%;  Score 19.4;  DB 2;  Length 224064;
  Best Local Similarity   95.2%;  Pred. No. 16;
  Matches   20;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTCCT 21
              |||||||||||||||||| |||
Db      31626 TGAGAGAGTTCTGGGTGGCCT 31606


RESULT 12
AY058623
LOCUS       AY058623                3521 bp    mRNA    linear   INV 17-OCT-2001
DEFINITION  Drosophila melanogaster LD31543 full length cDNA.
ACCESSION   AY058623
VERSION     AY058623.1  GI:16198106
KEYWORDS    FLI_CDNA.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 3521)
  AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Agbayani,A., Carlson,J.,
            Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George,R.,
            Gonzalez,M., Guarin,H., Li,P., Liao,G., Miranda,A., Mungall,C.J.,
            Nunoo,J., Pacleb,J., Paragas,V., Park,S., Phouanenavong,S., Wan,K.,
            Yu,C., Lewis,S.E., Rubin,G.M. and Celniker,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-OCT-2001) Berkeley Drosophila Genome Project,
            Lawrence Berkeley National Laboratory, One Cyclotron Road,
            Berkeley, CA 94720, USA
COMMENT     Sequence submitted by:
            Berkeley Drosophila Genome Project
            Lawrence Berkeley National Laboratory
            Berkeley, CA 94720
            This clone was sequenced as part of a high-throughput process to
            sequence clones from Drosophila Gene Collection 1 (Rubin et al.,
            Science 2000). The sequence has been subjected to integrity checks
            for sequence accuracy, presence of a polyA tail and contiguity
            within 100 kb in the genome. Thus we believe the sequence to
            reflect accurately this particular cDNA clone.  However, there are
            artifacts associated with the generation of cDNA clones that may
            have not been detected in our initial analyses such as internal
            priming, priming from contaminating genomic DNA, retained introns
            due to reverse transcription of unspliced precursor RNAs, and
            reverse transcriptase errors that result in single base changes.
            For further information about this sequence, including its location
            and relationship to other sequences, please visit our Web site
            (http://fruitfly.berkeley.edu) or send email to
            cdna@fruitfly.berkeley.edu.
FEATURES             Location/Qualifiers
     source          1. .3521
                     /organism="Drosophila melanogaster"
                     /mol_type="mRNA"
                     /strain="y; cn bw sp"
                     /db_xref="taxon:7227"
                     /map="18C8-18C8"
     gene            1. .3521
                     /gene="mRpS14"
                     /note="alignment with genomic scaffold AE003512. gene does
                     not completely overlap longest ORF"
                     /db_xref="FLYBASE:FBgn0044030"
     CDS             56. .3475
                     /gene="mRpS14"
                     /note="Longest ORF"
                     /codon_start=1
                     /product="LD31543p"
                     /protein_id="AAL13852.1"
                     /db_xref="GI:16198107"
                     /db_xref="FLYBASE:FBgn0044030"
                     /translation="MSRIKKKADEHASSRRRTQINLIEFSFLDHKAELVGLLQAKSRQ
                     VGRSLVENEADFWKFVGKYEAMMRSTGQPVLPPRLAEAELSSMQPYHRSKHLALRLDA
                     DELLQEARDQMQSQFRHILLVYLDFRQTEKFQRIRKLRQTQRNLPIARFRKDLREALD
                     TSRVVIVAGDTGCGKSTQVPQYLYDFGYRSIACTQPRRLACVSLCKRVAHELLDDYGS
                     RVAFQIRFERSRTKLTNILFITEGLLLRQLAVAANLDQYDALILDEIHERNLFGDFLL
                     GVTKCLLRARPQLKLILMSATINVELFHGYFGEEGARLVQVPGRLFPIKLRYLPPPAL
                     ELKAGQATSKRSQRNRIDPAPFVQVLSLIDQQYPTSERGDVLIFVSGVNEIESVVEAV
                     HEYATEQTHWLVLPLHSGQAIADQSKVFDYAPEGMRKCIVSTNIAETSLTVDGVRFVV
                     DSGKVKEMNFDATCKGQRLKEFWVSKSSADQRKGRAGRTGPGVCFRLYTAEQYNAFEA
                     YPTPEIYRVPLDTMLLQMVSMGLPDVRAFPFIEAPETERIEQTILALKQHCALSVEEK
                     ITPLGRSLANLPVELSIGKMLLMGSVFPEVEQLLTLAAMLSVQNPLTNRAHTDQRCVR
                     ERESLESDHGDLFTLVRLYNVWVQLKMRRDGTRQWCRRLGIEEQRFYEVTKLRQQFQR
                     ILESCGMVVASDSDSQLTSAERATRHGELRQLKAMKRRQRFEQPRQRKLLKQSAGRVA
                     EDEEEQEEAQGDDMRDVDFRLRFDPRQLALLERSSRLDRHSVVVLLKLLLGSGFYPQL
                     AISDEFNYCKGGGQQFFHTRLKPFVLQHPNSQFAKHFELLKLTESDLLPKPDFYTPKL
                     PLSKRHQLLCYQSLLETAKPYLINCIRLPAAQTLLLFSFAIDTNAGITQIACDGWLGL
                     DLPMPGSGMELLRRAIELRRRWSRLLYDKLDDLKSKQESPSHSSDDRSSALWQDLVDY
                     MALDVAYAIRRLLPADIKRLYTHQAPSARLAELKENPFAVDFPMTPNEEKGGLNVSEH
                     VVYSCLAEQQWTTAMDATLRAEPWQCTRCDFELEEFDVLEQLVHRSKCKRRKAQESRT
                     AKTTSTESSTEVASSSSSHSSGGYYCNSCKRELRLTTIDILRHKKQCRNNK"
ORIGIN

  Query Match             85.5%;  Score 18.8;  DB 3;  Length 3521;
  Best Local Similarity   90.9%;  Pred. No. 27;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTCCTA 22
              ||| |||||||||||||||| |
Db       1461 TGAAAGAGTTCTGGGTGTCCAA 1482


RESULT 13
AC015346/c
LOCUS       AC015346               26121 bp    DNA     linear   HTG 16-NOV-1999
DEFINITION  Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***.
ACCESSION   AC015346
VERSION     AC015346.1  GI:6435989
KEYWORDS    HTG; HTGS_PHASE2.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 26121)
  AUTHORS   Adams,M. and Venter,J.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-NOV-1999) Celera Genomics, 45 West Gude Drive,
            Rockville, MD, USA
COMMENT     This sequence was identified as CDM:10210169 by the submitter.
            For further information on this sequence e-mail to fly@celera.com.
            * NOTE: This is a 'working draft' sequence.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
FEATURES             Location/Qualifiers
     source          1. .26121
```

```
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
ORIGIN

  Query Match             85.5%;  Score 18.8;  DB 2;  Length 26121;
  Best Local Similarity   90.9%;  Pred. No. 31;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTCCTA 22
              ||| |||||||||||||||| |
Db       5183 TGAAAGAGTTCTGGGTGTCCAA 5162


RESULT 14
AC025824/c
LOCUS       AC025824               177917 bp    DNA     linear   PRI 19-MAR-2002
DEFINITION  Homo sapiens chromosome 15 clone RP11-430M3, complete sequence.
ACCESSION   AC025824
VERSION     AC025824.37  GI:19549230
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 177917)
  AUTHORS   Abola,A.P., Bruno,D., Conn,L., Dela Rosa,M., Faulkner,D.,
            Federspiel,N., Glukhov,S., Hansen,N., Herman,Z.S., Hyman,R.,
            Mao,J., Komp,C., Kottler,S., Lam,B., Marathe,R., Miranda,M.,
            Morehouse,A.J., Nguyen,M., Oefner,P., Palm,C.J., Ramirez,D.,
            Southwick,A.M., Webb,C., Wilhelmy,J., Yu,S. and Davis,R.W.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 177917)
  AUTHORS   Bruno,D., Conn,L., Dela Rosa,M., Faulkner,D., Federspiel,N.,
            Glukhov,S., Hansen,N., Hyman,R., Mao,J., Marathe,R.,
            Morehouse,A.J., Oefner,P., Palm,C.J., Ramirez,D., Wilhelmy,J.,
            Yu,S. and Davis,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-MAR-2000) DNA Sequencing and Technology Center,
            Stanford University, 855 California Avenue, Palo Alto, CA 94304,
            USA
REFERENCE   3  (bases 1 to 177917)
  AUTHORS   Abola,A.P., Bruno,D., Conn,L., Dela Rosa,M., Faulkner,D.,
            Federspiel,N., Glukhov,S., Hansen,N., Herman,Z.S., Hyman,R.,
            Komp,C., Kottler,S., Lam,B., Mao,J., Marathe,R., Miranda,M.,
            Morehouse,A.J., Nguyen,M., Oefner,P., Palm,C.J., Ramirez,D.,
            Southwick,A.M., Webb,C., Wilhelmy,J., Yu,S. and Davis,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-MAR-2002) Genome Technology Center, Stanford
            University, 855 California Avenue, Palo Alto, CA 94304, USA
COMMENT     On Mar 19, 2002 this sequence version replaced gi:18252013.
            -------------- Genome Center
                 Center: Stanford DNA Sequencing and Technology Development
            Center
                 Center code: SDSTDC
                 Web site: http://sequence-www.stanford.edu/group/human/
                 Contact: hum-info@sequence.stanford.edu
            -------------- Project Information
                 Center project name: 857
                 Center clone name: RP11-430M3

            -------------- Summary
            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were  covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.
FEATURES              Location/Qualifiers
     source           1. .177917
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /chromosome="15"
                      /clone="RP11-430M3"
                      /clone_lib="RPCI human BAC library 11"
     misc_feature     128. .131
                      /note="Target site of Tn10 insertion in BAC.  Tn10
                      sequence and target site duplication, CATT, have been
                      removed from completed sequence."
     misc_feature     23502. .23850
                      /note="Single-stranded, single chemistry with consensus
                      quality below 30."
ORIGIN

  Query Match             85.5%;  Score 18.8;  DB 9;  Length 177917;
  Best Local Similarity   90.9%;  Pred. No. 35;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTCCTA 22
              | |||||||||||||| ||||||
Db     151626 TCAGAGAGTTCTGGGGGTCCTA 151605


RESULT 15
AC010847/c
LOCUS       AC010847               180213 bp    DNA     linear   INV 17-MAR-2001
DEFINITION  Drosophila melanogaster, chromosome X, region 18D-18D, BAC clone
            BACR10M08, complete sequence.
ACCESSION   AC010847
VERSION     AC010847.11  GI:13374650
KEYWORDS    HTG.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 180213)
  AUTHORS   Celniker,S.E., Adams,M.D., Kronmiller,B., Tyler,D., Wan,K.H.,
            Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Brandon,R.C.,
            Rogers,Y., An,H., Baldwin,D., Banzon,J., Beeson,K.Y., Busam,D.A.,
            Carlson,J.W., Center,A., Champe,M., Davenport,L.B., Dietz,S.M.,
            Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D.,
            Ferriera,S., Frise,E., Galle,R.F., Garg,N.S., George,R.A.,
            Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J.,
            Ibegwam,C., Jalali,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A.,
            McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J.,
            Pacleb,J., Paragas,V., Park,S., Patel,S., Pfeiffer,B.,
            Phouanenavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F.,
            Stapleton,M., Strong,R., Svirskas,R., Tector,C., Williams,S.M.,
            Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.
  TITLE     Sequencing of Drosophila chromosome X, region 18D-18D
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 180213)
  AUTHORS   Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
            Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
            Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
            Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
            Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,
            Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
            Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E.,
            Svirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and
            Rubin,G.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-SEP-1999) Drosophila Genome Center, Lawrence Berkeley
            Laboratory, MS 64-121, Berkeley, CA 94720, USA
COMMENT     On Mar 17, 2001 this sequence version replaced gi:6563418.
            Sequence submitted by:
            Berkeley Drosophila Genome Project
            Lawrence Berkeley National Laboratory, MS 64-121
            Berkeley, CA 94720
            This sequence was assembled using end sequences from a whole genome
            shotgun and from subclones of this BAC and its neighboring clones.
            For further information about this sequence, including its location
            and relationship to other sequences, please visit our sequence
```

archive Web site (http://www.fruitfly.org/sequence/) or send email to bdgp@fruitfly.berkeley.edu.

```
FEATURES             Location/Qualifiers
     source          1. .180213
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /strain="y; cn bw sp"
                     /db_xref="taxon:7227"
                     /chromosome="X"
                     /map="18D-18D"
                     /clone="BACR10M08 (D904)"
                     /clone_lib="RPCI-98 (Roswell Park Cancer Institute
                     Drosophila melanogaster BAC library, partial EcoRI in
                     pBACe3.6)"
ORIGIN
```

Query Match 85.5%; Score 18.8; DB 3; Length 180213;
Best Local Similarity 90.9%; Pred. No. 35;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 TGAGAGAGTTCTGGGTGTCCTA 22
              ||| |||||||||||||||| |
Db      18888 TGAAAGAGTTCTGGGTGTCCAA 18867
```

Search completed: February 19, 2004, 23:23:28
Job time : 101.318 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 21:15:48 ; Search time 20.3821 Seconds
(without alignments)
4585.415 Million cell updates/sec

Title: US-09-939-853A-142
Perfect score: 22
Sequence: 1 tgagagagttctgggtgtccta 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
	1	22	100.0	22	6	ABK61507	Abk61507 Human NOV
	2	22	100.0	1183	6	ABK61465	Abk61465 Human cDN
	3	22	100.0	2567	6	AAD43980	Aad43980 Human Src
	4	20.4	92.7	763	6	ABQ98670	Abq98670 Human ORF
	5	18.8	85.5	3319	4	ABL28809	Abl28809 Drosophil
c	6	18.8	85.5	5782	4	ABL28808	Abl28808 Drosophil
	7	17.4	79.1	2442	7	ABX70675	Abx70675 Human cDN
	8	17.4	79.1	2789	6	AAD30568	Aad30568 Human pro
	9	17.4	79.1	3076	4	AAH57383	Aah57383 Human ske
	10	17.4	79.1	3210	5	AAS90933	Aas90933 DNA encod
	11	17.4	79.1	3327	8	ACH03882	Ach03882 Human cDN
	12	17.4	79.1	5149	2	AAT32455	Aat32455 Calpain 1
c	13	17.2	78.2	355	4	AAK67913	Aak67913 Human imm
c	14	17.2	78.2	356	4	AAK67912	Aak67912 Human imm
c	15	17.2	78.2	359	4	AAK56372	Aak56372 Human imm
	16	17.2	78.2	1163	7	ACA49310	Aca49310 Prokaryot
	17	17.2	78.2	1182	7	ACA51118	Aca51118 Prokaryot
	18	17.2	78.2	1182	7	ACA51577	Aca51577 Prokaryot
	19	17.2	78.2	1185	7	ACA35390	Aca35390 Prokaryot
c	20	17.2	78.2	1457	5	AAC93029	Aac93029 Viral rep
	21	17.2	78.2	13327	4	ABL02914	Abl02914 Drosophil
	22	16.8	76.4	340	7	ADA72333	Ada72333 Rice gene
	23	16.8	76.4	1382	4	AAK73700	Aak73700 Human imm
c	24	16.8	76.4	2394	9	ADB62889	Adb62889 Human cDN
c	25	16.8	76.4	5313	4	ABL06558	Abl06558 Drosophil
	26	16.8	76.4	28772	6	ABK83555	Abk83555 Human cDN
	27	16.4	74.5	701	4	AAF23001	Aaf23001 Human pro
	28	16.4	74.5	1191	6	ABL90353	Abl90353 Human pol
	29	16.4	74.5	1273	2	AAT64550	Aat64550 Activated
	30	16.4	74.5	1357	5	AAH68577	Aah68577 Human pro
	31	16.4	74.5	1373	2	AAT64549	Aat64549 Human act
	32	16.4	74.5	1621	2	AAZ20855	Aaz20855 Polynucle
	33	16.4	74.5	1621	4	AAS59292	Aas59292 Human cDN
	34	16.4	74.5	1621	6	ABA90961	Aba90961 Human pol
	35	16.4	74.5	2275	4	ABL19829	Abl19829 Drosophil
	36	16.4	74.5	2857	4	ABL13023	Abl13023 Drosophil
c	37	16.4	74.5	2872	4	AAI61360	Aai61360 Human pol
c	38	16.4	74.5	3128	6	ABZ35472	Abz35472 Human gen
c	39	16.4	74.5	3128	10	ADE77208	Ade77208 Human cDN
c	40	16.4	74.5	3986	5	ABV28780	Abv28780 Human pro
c	41	16.4	74.5	3986	5	ABV22948	Abv22948 Human pro
	42	16.4	74.5	4665	6	ABK83947	Abk83947 Human cDN
c	43	16.4	74.5	5912	4	ABL19828	Abl19828 Drosophil
c	44	16.4	74.5	7420	4	ABL13022	Abl13022 Drosophil
	45	16.4	74.5	27684	8	ADA03062	Ada03062 Human KCN

ALIGNMENTS

RESULT 1
ABK61507
ID ABK61507 standard; DNA; 22 BP.
XX
AC ABK61507;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human NOV13 RT-PCR primer #2.
XX
KW Human; ss; NOVX; gene therapy; cardiomyopathy; atherosclerosis;
KW cell signal processing disorder; metabolic pathway modulation disorder;
KW diabetes; cancer; adenocarcinoma; lymphoma; prostate cancer; primer;
KW uterus cancer; immune response; graft-versus-host disease; Exon linking;
KW acquired immunodeficiency syndrome; AIDS; asthma; Crohn's disease;
KW hypertension; congenital heart defects; multiple sclerosis; inflammation;
KW Albright hereditary osteodystrophy; reverse transcriptase PCR.
XX
OS Homo sapiens.
XX
PN WO200216599-A2.
XX
PD 28-FEB-2002.
XX
PF 27-AUG-2001; 2001WO-US026510.
XX
PR 25-AUG-2000; 2000US-0228191P.
PR 08-FEB-2001; 2001US-0267300P.
PR 20-FEB-2001; 2001US-0269961P.
PR 20-MAR-2001; 2001US-0277337P.
XX
PA (CURA-) CURAGEN CORP.
PA (CORT-) COR THERAPEUTICS INC.
XX
PI Burgess CE, Conley PB, Grosse WM, Hart M, Kekuda R, Shimkets RA;
PI Spytek KA, Szekeres ES, Tomlinson JE, Topper JN, Yang R;
XX
DR WPI; 2002-280937/32.
XX
PT New polypeptides for treating or preventing a disorder associated with
PT them, in humans, e.g. cardiomyopathy, atherosclerosis or cancers.
XX
PS Example 2; Page 234; 263pp; English.
XX
CC The invention relates to an isolated polypeptide (NOVX) a mature form of
CC NOVX, a NOVX variant (differing by no more than 15%), the nucleotide

```
CC   encoding NOVX (or its complement, fragment or variant). NOVX is NOV1-14,
CC   15a, 15b, 16a, and 16b. The NOVX polypeptide, nucleic acid encoding it
CC   and antibody against it, are useful for treating or preventing (e.g. by
CC   gene therapy) a NOVX-associated disorder in humans, e.g. cardiomyopathy,
CC   atherosclerosis, a disorder related to cell signal processing and
CC   metabolic pathway modulation, diabetes or cancers. The NOVX polypeptide
CC   and nucleic acids are also useful for determining the presence of
CC   predisposition to the diseases. The NOVX nucleic acid and polypeptide are
CC   especially useful in therapeutic or prophylactic applications for
CC   disorders associated with aberrant NOVX expression or activity, e.g.
CC   cancers (e.g. adenocarcinoma, lymphoma, prostate cancer or uterus
CC   cancer), immune response, graft-versus-host disease, acquired
CC   immunodeficiency syndrome (AIDS), asthma, Crohn's disease, hypertension,
CC   congenital heart defects, multiple sclerosis, inflammation or Albright
CC   hereditary osteodystrophy and many other diseases listed in the
CC   specification. The DNA encoding the protein is useful in gene therapy for
CC   treating the conditions. This is also useful in detection assays,
CC   chromosome mapping, tissue typing, diagnostic or prognostic assays, or
CC   for developing a powerful assay system for functional analysis of various
CC   human disorders, as well as in diagnostic applications. The present
CC   sequence is a reverse transcriptase (RT)-PCR primer used to measure
CC   tissue specific expression of mRNA encoding a NOVX protein
XX
SQ   Sequence 22 BP; 4 A; 3 C; 8 G; 7 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 22;  DB 6;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 0.79;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTCCTA 22
              ||||||||||||||||||||||
Db          1 TGAGAGAGTTCTGGGTGTCCTA 22


RESULT 2
ABK61465
ID   ABK61465 standard; cDNA; 1183 BP.
XX
AC   ABK61465;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Human cDNA encoding protein NOV13.
XX
KW   Human; gene; ss; NOVX; gene therapy; cardiomyopathy; atherosclerosis;
KW   cell signal processing disorder; metabolic pathway modulation disorder;
KW   diabetes; cancer; adenocarcinoma; lymphoma; prostate cancer;
KW   uterus cancer; immune response; graft-versus-host disease;
KW   acquired immunodeficiency syndrome; AIDS; asthma; Crohn's disease;
KW   hypertension; congenital heart defects; multiple sclerosis; inflammation;
KW   Albright hereditary osteodystrophy.
XX
OS   Homo sapiens.
XX
PN   WO200216599-A2.
XX
PD   28-FEB-2002.
XX
PF   27-AUG-2001; 2001WO-US026510.
XX
PR   25-AUG-2000; 2000US-0228191P.
PR   08-FEB-2001; 2001US-0267300P.
PR   20-FEB-2001; 2001US-0269961P.
PR   20-MAR-2001; 2001US-0277337P.
XX
PA   (CURA-) CURAGEN CORP.
PA   (CORT-) COR THERAPEUTICS INC.
XX
PI   Burgess CE,  Conley PB,  Grosse WM,  Hart M,  Kekuda R,  Shimkets RA;
PI   Spytek KA,  Szekeres ES,  Tomlinson JE,  Topper JN,  Yang R;
XX
DR   WPI; 2002-280937/32.
DR   P-PSDB; AAU91308.
XX
PT   New polypeptides for treating or preventing a disorder associated with
PT   them, in humans, e.g. cardiomyopathy, atherosclerosis or cancers.
XX
PS   Claim 1; Page 98; 263pp; English.
XX
CC   The invention relates to an isolated polypeptide (NOVX) a mature form of
CC   NOVX, a NOVX variant (differing by no more than 15%), the nucleotide
CC   encoding NOVX (or its complement, fragment or variant). NOVX is NOV1-14,
CC   15a, 15b, 16a, and 16b. The NOVX polypeptide, nucleic acid encoding it
CC   and antibody against it, are useful for treating or preventing (e.g. by
CC   gene therapy) a NOVX-associated disorder in humans, e.g. cardiomyopathy,
CC   atherosclerosis, a disorder related to cell signal processing and
CC   metabolic pathway modulation, diabetes or cancers. The NOVX polypeptide
CC   and nucleic acids are also useful for determining the presence of
CC   predisposition to the diseases. The NOVX nucleic acid and polypeptide are
CC   especially useful in therapeutic or prophylactic applications for
CC   disorders associated with aberrant NOVX expression or activity, e.g.
CC   cancers (e.g. adenocarcinoma, lymphoma, prostate cancer or uterus
CC   cancer), immune response, graft-versus-host disease, acquired
CC   immunodeficiency syndrome (AIDS), asthma, Crohn's disease, hypertension,
CC   congenital heart defects, multiple sclerosis, inflammation or Albright
CC   hereditary osteodystrophy and many other diseases listed in the
CC   specification. The DNA encoding the protein is useful in gene therapy for
CC   treating the conditions. This is also useful in detection assays,
CC   chromosome mapping, tissue typing, diagnostic or prognostic assays, or
CC   for developing a powerful assay system for functional analysis of various
CC   human disorders, as well as in diagnostic applications. The present
CC   sequence encodes a NOVX protein
XX
SQ   Sequence 1183 BP; 251 A; 359 C; 333 G; 240 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 22;  DB 6;  Length 1183;
  Best Local Similarity   100.0%;  Pred. No. 1.2;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTCCTA 22
              ||||||||||||||||||||||
Db        224 TGAGAGAGTTCTGGGTGTCCTA 245


RESULT 3
AAD43980
ID   AAD43980 standard; cDNA; 2567 BP.
XX
AC   AAD43980;
XX
DT   13-DEC-2002  (first entry)
XX
DE   Human Src-Like Adapter Protein-2 (hSLAP-2) cDNA.
XX
KW   Human; SH2/SH3-domain-containing adapter; Src-Like Adapter Protein-2;
KW   SLAP-2; immune disorder; signal transduction; autoimmune disease; cancer;
KW   neoplasm; inflammation disorder; rheumatoid arthritis; osteoarthritis;
KW   psoriasis; rhinitis; allergy; inflammatory bowel disease; gene therapy;
KW   Crohn's disease; systemic lupus erythematosus; tissue/organ rejection;
KW   multiple sclerosis; asthma; acute respiratory distress syndrome;
KW   pulmonary disorder; dermatological; neuroprotective; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             415. .1200
FT                   /*tag=  a
FT                   /product= "Human SLAP-2"
XX
PN   WO200242457-A1.
XX
PD   30-MAY-2002.
XX
PF   20-NOV-2001; 2001WO-US043367.
XX
PR   22-NOV-2000; 2000US-0252545P.
```

XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Chang H, Yang W, Wu Y, Whitney GS, Perez-Villar JJ, Kanner SB;
XX
DR WPI; 2002-463632/49.
DR P-PSDB; AAE26357.
XX
PT Novel substantially purified human SH2/SH3-domain-containing adapter
PT polypeptide, termed Src-Like Adapter Protein-2, useful for therapeutic
PT intervention in immunological and inflammatory disorders and cancer.
XX
PS Claim 2; Fig 1; 85pp; English.
XX
CC The invention relates to a substantially purified human SH2/SH3-domain-
CC containing adapter polypeptide, termed Src-Like Adapter Protein-2 (SLAP-
CC 2). The invention is useful for treating an immune disorder involving
CC hyperactivity of B- or T- lymphocytes in a mammal. The invention is
CC useful for screening for antagonists or inhibitors of the interaction of
CC hSLAP-2 with cellular signalling compounds, for diagnosing, treating or
CC preventing diseases or disorders associated with aberrant or uncontrolled
CC cellular signal transduction, for determining those cellular signalling
CC molecules which associate with hSLAP-2 and which provide critical signals
CC for cell activation, and as effectors in methods to affect T- cell
CC activation. The invention is useful in screening assays to identity and
CC detect candidate bioactive agents that modulate hSLAP-2 bioactivity, for
CC potential use to treat autoimmune diseases which may be caused by
CC hyperactivated B cells, as well as to treat diseases which may be caused
CC by hyperactivated T cells, in addition to other immune system related
CC conditions, diseases, or disorders, T-cell and B-cell neoplasms,
CC inflammation disorders, diseases and conditions, rheumatoid arthritis,
CC osteoarthritis, psoriasis, rhinitis, inflammatory bowel disease (Crohn's
CC and ulcerative colitis), allergies, particularly those involving
CC hyperactivity of B-cells and T- cells, or other immune cells, such as
CC mast cells or eosinophils, autoimmune diseases such as systemic lupus
CC erythematosus and multiple sclerosis, pulmonary diseases including
CC asthma, acute respiratory distress syndrome, and chronic obstructive
CC pulmonary disorder, tissue/ organ rejection and cancer. The invention is
CC useful in gene therapy. The present sequence is human SLAP-2 cDNA
XX
SQ Sequence 2567 BP; 611 A; 741 C; 666 G; 549 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 6; Length 2567;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGAGAGAGTTCTGGGTGTCCTA 22
||||||||||||||||||||||
Db 241 TGAGAGAGTTCTGGGTGTCCTA 262

RESULT 4
ABQ98670
ID ABQ98670 standard; DNA; 763 BP.
XX
AC ABQ98670;
XX
DT 04-NOV-2002 (first entry)
XX
DE Human ORF477 coding sequence.
XX
KW Cytostatic; Cardiant; Anti-allergic; Immunosuppressive; Vulnerary;
KW Antiinflammatory; gene therapy; human; ORFX; atherogenic; platelet;
KW human umbilical vein endothelial cell; HUVEC; atherosclerotic plaque;
KW cancer; cardiovascular disease; allergy; autoimmune disease;
KW wound healing; blood coagulation disorder; inflammatory disorder; ds.
XX
OS Homo sapiens.
XX
PN US2002082206-A1.
XX
PD 27-JUN-2002.
XX
PF 30-MAY-2001; 2001US-00867550.
XX
PR 30-MAY-2000; 2000US-0208427P.
XX
PA (LEAC/) LEACH M D.
PA (MEHR/) MEHRABAN F.
PA (CONL/) CONLEY P B.
PA (TOPP/) TOPPER J N.
PA (LAWD/) LAW D.
XX
PI Leach MD, Mehraban F, Conley PB, Topper JN, Law D;
XX
DR WPI; 2002-626554/67.
DR P-PSDB; ABP64107.
XX
PT New polypeptide designated ORFX are present in human atherogenic cells
PT and are useful to prevent and treat ORFX-associated disorders including
PT cancer, allergy, wound healing or autoimmune, cardiovascular or
PT inflammatory disease.
XX
PS Claim 2; SEQ ID NO 953; 78pp; English.
XX
CC The present invention relates to novel human ORFX polypeptides and their
CC coding sequences (ABP63631-ABP64681 and ABQ98194-ABQ99267). The sequences
CC were discovered in human atherogenic cells, in particular in platelets
CC and human umbilical vein endothelial cells (HUVEC) and are expressed in
CC many other tissues as well. Atherogenic cells are cells which have the
CC potential to develop atherosclerotic plaques. The ORFX polypeptides and
CC nucleic acids are useful for treating or preventing a pathological
CC condition associated with an ORFX-associated disorder, e.g. cancer,
CC cardiovascular disease, allergy, autoimmune disease, wound healing, blood
CC coagulation disorders or inflammatory disorders. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from the USPTO web site at
CC seqdata.uspto.gov/sequence.html?DocID=20020082206
XX
SQ Sequence 763 BP; 176 A; 222 C; 218 G; 147 T; 0 U; 0 Other;

Query Match 92.7%; Score 20.4; DB 6; Length 763;
Best Local Similarity 95.5%; Pred. No. 6.7;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TGAGAGAGTTCTGGGTGTCCTA 22
||||||||||||||| |||||||
Db 112 TGAGAGAGTTCTGGATGTCCTA 133

RESULT 5
ABL28809
ID ABL28809 standard; DNA; 3319 BP.
XX
AC ABL28809;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 37900.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX

PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Claim 1; SEQ ID NO 37900; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 3319 BP; 699 A; 947 C; 992 G; 681 T; 0 U; 0 Other;

Query Match 85.5%; Score 18.8; DB 4; Length 3319;
Best Local Similarity 90.9%; Pred. No. 45;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TGAGAGAGTTCTGGGTGTCCTA 22
||| |||||||||||||||| |
Db 1729 TGAAAGAGTTCTGGGTGTCCAA 1750

RESULT 6
ABL28808/c
ID ABL28808 standard; DNA; 5782 BP.
XX
AC ABL28808;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 37897.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Claim 1; SEQ ID NO 37897; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 5782 BP; 1326 A; 1604 C; 1561 G; 1291 T; 0 U; 0 Other;

Query Match 85.5%; Score 18.8; DB 4; Length 5782;
Best Local Similarity 90.9%; Pred. No. 48;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TGAGAGAGTTCTGGGTGTCCTA 22
||| |||||||||||||||| |
Db 2789 TGAAAGAGTTCTGGGTGTCCAA 2768

RESULT 7
ABX70675
ID ABX70675 standard; cDNA; 2442 BP.
XX
AC ABX70675;
XX
DT 04-MAR-2003 (first entry)
XX
DE Human cDNA encoding NOV26.
XX
KW Human; ss; gene; NOVX; metabolic disorder; diabetes; cardiomyopathy;
KW obesity; infectious disease; anorexia; neurodegenerative disorder;
KW Alzheimer's disease; Parkinson's disease; immune disorder;
KW haematopoietic disorder; dyslipidaemia; metabolic disturbance;
KW metabolic syndrome X; wasting disorder; cancer; gene therapy; SNP;
KW single nucleotide polymorphism.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT variation replace(72,C)
FT /*tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(78,C)
FT /*tag= b
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(201,C)
FT /*tag= c
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(412,T)
FT /*tag= d
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(2253,G)
FT /*tag= e
FT /standard_name= "Single nucleotide polymorphism"
XX
PN WO200281518-A2.
XX
PD 17-OCT-2002.
XX
PF 21-FEB-2002; 2002WO-US005374.
XX
PR 21-FEB-2001; 2001US-0270220P.
PR 21-FEB-2001; 2001US-0270523P.
PR 23-FEB-2001; 2001US-0270797P.
PR 23-FEB-2001; 2001US-0270810P.
PR 08-MAR-2001; 2001US-0274295P.
PR 16-MAR-2001; 2001US-0276400P.
PR 16-MAR-2001; 2001US-0276677P.
PR 26-MAR-2001; 2001US-0278796P.
PR 04-APR-2001; 2001US-0281521P.
PR 25-APR-2001; 2001US-0286548P.
PR 17-MAY-2001; 2001US-0291765P.
PR 10-AUG-2001; 2001US-0311595P.
PR 13-AUG-2001; 2001US-0311980P.
PR 10-SEP-2001; 2001US-0318526P.

PR 17-SEP-2001; 2001US-0322712P.
PR 18-OCT-2001; 2001US-0330307P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Pena CEA, Shimkets RA, Li L, Shenoy SG, Kekuda R, Spytek KA;
PI Vernet CAM, Malyankar UL, Guo X, Gusev VY, Casman SJ, Boldog FL;
PI Furtak K, Tchernev VT, Patturajan M, Gangolli EA, Padigaru M, Liu X;
PI Baumgartner JC, Gerlach VL, Spaderna SK, Zerhusen BD;
XX
DR WPI; 2003-046859/04.
DR P-PSDB; ABU52620.
XX
PT New isolated NOVX polypeptide useful for treating cardiomyopathy,
PT atherosclerosis, metabolic disorders, diabetes, obesity, infectious
PT disease, anorexia, neurodegenerative disorders, Alzheimer's disease and
PT cancer.
XX
PS Claim 9; Page 210-211; 479pp; English.
XX
CC The invention relates to an isolated polypeptide termed NOVX (NOV1, 2a,
CC 2b, 3a, 3b, 4a, 4b, 5, 6, 7a-c, 8a-e, 9a-b, 10, 11, 12a-c, 13, 14, 15,
CC 16a-d, 17a-b, 18, 19, 20a-b, 21-30) appearing as ABU52578-ABU52624), a
CC variant of NOVX, a mature form of NOVX, and a variant of the mature form
CC of NOVX. Also included are a nucleic acid molecule (NOVX NA) encoding
CC NOVX, or a fragment or complement of NOVX NA, a vector comprising NOVX
CC NA, a cell comprising the vector, an anti-NOVX antibody (ab), determining
CC the presence or amount of NOVX or NOVX NA in a sample, and identifying an
CC agent that binds or modulates the expression or activity of NOVX. NOVX,
CC NOVX NA or ab is useful for treating or preventing a NOVX-associated
CC disorder in a subject, preferably human. Ab is useful for determining the
CC presence or amount of NOVX in a sample. NOVX is useful for identifying an
CC agent that binds to NOVX . NOVX, NOVX NA or ab is useful for treating
CC metabolic disorders, diabetes, cardiomyopathy, obesity, infectious
CC disease, anorexia, neurodegenerative disorders, Alzheimer's disease,
CC Parkinson's disease, immune disorders, haematopoietic disorders, and
CC various dyslipidaemias, metabolic disturbances associated with obesity,
CC the metabolic syndrome X and wasting disorders associated with chronic
CC diseases, various cancers, endocrine, connective tissue, blood, vascular,
CC skin, renal, bone, brain, muscle disorders, or bacterial, fungal,
CC protozoal or viral infections. NOVX, NOVX NA or ab is useful in screening
CC assays, detection assays, predictive medicine, and in methods of
CC treatment. NOVX is useful as immunogen, to screen for potential
CC ant/agonist compounds, and as bait protein in a two-hybrid or three-
CC hybrid assay. NOVX NA is useful in gene therapy, to express NOVX, to
CC detect NOVX mRNA or a genetic lesion in a NOVX gene, and to modulate NOVX
CC activity. The cell is useful for producing non-human transgenic animals.
CC Ab is useful for isolating, and purifying NOVX and to monitor protein
CC levels in tissue as part of a clinical testing procedure. The present
CC sequence encodes a NOVX protein
XX
SQ Sequence 2442 BP; 645 A; 631 C; 643 G; 523 T; 0 U; 0 Other;

Query Match 79.1%; Score 17.4; DB 7; Length 2442;
Best Local Similarity 94.7%; Pred. No. 2e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GAGAGTTCTGGGTGTCCTA 22
||||||||||| |||||||
Db 1116 GAGAGTTCTGGATGTCCTA 1134

RESULT 8
AAD30568
ID AAD30568 standard; cDNA; 2789 BP.
XX
AC AAD30568;
XX
DT 21-MAY-2002 (first entry)
XX
DE Human protease, PRTS-1 cDNA.
XX
KW Human; protease; PRTS-1; enzyme; gastritis; cirrhosis; Crohn's disease;
KW gastrointestinal disorder; autoimmune; inflammatory; cell proliferative;
KW cardiovascular; developmental; epithelial; neurological; reproductive;
KW AIDS; Acquired Immune Deficiency Syndrome; allergy; rheumatoid arthritis;
KW anaemia; asthma; atherosclerosis; hypertension; myocardial infarction;
KW hepatitis; cancer; psoriasis; Cushing's syndrome; hypothyroidism; eczema;
KW epilepsy; Alzheimer's disease; Huntington's disease; Parkinson's disease;
KW Pick's disease; infertility; vitiligo; drug screening; gene therapy; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 303. .2606
FT /*tag= a
FT /product= "Human PRTS-1 protein"
FT sig_peptide 303. .347
FT /*tag= b
FT mat_peptide 348. .2603
FT /*tag= c
FT /product= "Mature PRTS-1 protein"
XX
PN WO200208396-A2.
XX
PD 31-JAN-2002.
XX
PF 17-JUL-2001; 2001WO-US022397.
XX
PR 21-JUL-2000; 2000US-0220063P.
PR 28-JUL-2000; 2000US-0221680P.
PR 04-AUG-2000; 2000US-0223544P.
PR 11-AUG-2000; 2000US-0224717P.
PR 16-AUG-2000; 2000US-0225988P.
PR 23-AUG-2000; 2000US-0227568P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Delegeane AM, Gandhi AR, Hafalia AJA, Lu DAM, Patterson C;
PI Tribouley CM, Das D, Kallick DA, Nguyen DB, Lee EA, Khan FA, Yue H;
PI Au-Young J, Griffin JA, Policky JL, Ramkumar J, Yang J;
PI Thangavelu K, Ding L, Kearney L, Baughn MR, Borowsky ML;
PI Sanjanwala MS, Yao MG, Burford N, Walia NK, Lal P, Lee S, Todd S;
PI Lo TP, Tang YT, Elliott VS, Azimzai Y, Lu Y;
XX
DR WPI; 2002-206082/26.
DR P-PSDB; AAE19164.
XX
PT New human protease polypeptide, useful in diagnosis, prevention and
PT treatment of gastrointestinal, cardiovascular, autoimmune/inflammatory,
PT cell proliferative, developmental, epithelial and neurological disorders.
XX
PS Claim 5; Page 164-165; 182pp; English.
XX
CC The invention relates to an isolated human protease polypeptide (PRTS).
CC PRTS protein and DNA are useful for diagnosing, treating and preventing
CC gastrointestinal disorders (gastritis, cirrhosis, Crohn's disease),
CC autoimmune/inflammatory disorders (AIDS, allergy, rheumatoid arthritis,
CC anaemia, asthma), cardiovascular disorder (atherosclerosis, hypertension,
CC myocardial infarction), cell proliferative disorders (hepatitis, cancer,
CC psoriasis), developmental disorders (Cushing's syndrome, hypothyroidism),
CC epithelial disorder (vitiligo, keloid, eczema), neurological disorders
CC (epilepsy, Alzheimer's disease, Pick's disease, Huntington's disease,
CC Parkinson's disease), and reproductive disorders (infertility). PRTS
CC protein is useful in a number of drug screening techniques and to analyse
CC the proteome of a tissue or cell type. PRTS DNA is useful for creating
CC knockin humanised animals or transgenic animals to model human diseases,
CC in somatic or germline gene therapy and in microarrays utilising fluids
CC or tissues from patients to detect altered PKIN expression. The present
CC sequence is human PRTS-1 cDNA
XX
SQ Sequence 2789 BP; 715 A; 745 C; 698 G; 631 T; 0 U; 0 Other;

Query Match 79.1%; Score 17.4; DB 6; Length 2789;
Best Local Similarity 94.7%; Pred. No. 2.1e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          4 GAGAGTTCTGGGTGTCCTA 22
              ||||||||||| |||||||
Db       1342 GAGAGTTCTGGATGTCCTA 1360
```

RESULT 9
AAH57383
ID AAH57383 standard; cDNA; 3076 BP.
XX
AC AAH57383;
XX
DT 10-SEP-2001 (first entry)
XX
DE Human skeletal muscle cell specific cDNA sequence SEQ ID NO:223.
XX
KW Human; tissue specific; diagnosis; brain; heart; skeletal muscle; lung;
KW liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;
KW metabolic disease; developmental disease; cytostatic; immunomodulatory;
KW neuroprotective; gene therapy; cancer; immunopathology; neuropathology.
XX
OS Homo sapiens.
XX
PN WO200132927-A2.
XX
PD 10-MAY-2001.
XX
PF 02-NOV-2000; 2000WO-US030396.
XX
PR 04-NOV-1999; 99US-0163508P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Sornasse T, Seilhamer JJ, Watson GA;
XX
DR WPI; 2001-291057/30.
XX
PT New cell and tissue specific polynucleotides useful for diagnosis,
PT prognosis or monitoring of treatments for disorders where the gene is
PT associated with a cancer, immunopathology or neuropathology.
XX
PS Claim 1; Page 153-154; 327pp; English.
XX
CC AAH57161 to AAH57576 represent cell and tissue specific polynucleotide
CC sequences (I). (I) can have cytostatic, immunomodulatory and
CC neuroprotective activities, and can be used in gene therapy. (I) and
CC proteins (II) encoded by then are used in high throughput screening
CC assays to select DNA molecules, RNA molecules, peptide nucleic acids,
CC mimetics, peptides, proteins, agonists, antagonists, antibodies or their
CC fragments, immunoglobulins, inhibitors, drug compounds and pharmaceutical
CC agents. Expression of (I) in a sample indicates the differentiation of
CC embryonic stem cells into a tissue selected from brain, heart, kidney,
CC liver, lung, skeletal muscle or pancreatic tissues. (I) and (II) are used
CC to produce an expression profile that defines a metabolic or
CC developmental process, treatment, condition, disease or disorder. The
CC gene profile can be used for diagnosis, prognosis or monitoring of
CC treatments and for investigating a predisposition to a disorder where the
CC gene is associated with a cancer, immunopathology or neuropathology
XX
SQ Sequence 3076 BP; 800 A; 814 C; 792 G; 670 T; 0 U; 0 Other;

```
  Query Match             79.1%;  Score 17.4;  DB 4;  Length 3076;
  Best Local Similarity   94.7%;  Pred. No. 2.1e+02;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GAGAGTTCTGGGTGTCCTA 22
              ||||||||||| |||||||
Db       1244 GAGAGTTCTGGATGTCCTA 1262
```

RESULT 10
AAS90933
ID AAS90933 standard; cDNA; 3210 BP.
XX
AC AAS90933;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #26737.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG26746.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 26737; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 3210 BP; 821 A; 839 C; 861 G; 689 T; 0 U; 0 Other;

```
  Query Match             79.1%;  Score 17.4;  DB 5;  Length 3210;
  Best Local Similarity   94.7%;  Pred. No. 2.1e+02;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GAGAGTTCTGGGTGTCCTA 22
              ||||||||||| |||||||
Db       1784 GAGAGTTCTGGATGTCCTA 1802
```

RESULT 11
ACH03882
ID ACH03882 standard; cDNA; 3327 BP.
XX
AC ACH03882;

```
XX
DT   26-SEP-2003  (first entry)
XX
DE   Human cDNA differentially expressed in lung cancer #87.
XX
KW   Gene therapy; emphysema; ss; gene; chronic obstructive pulmonary disease;
KW   respiratory disorder; lung cancer; asthma; human.
XX
OS   Homo sapiens.
XX
PN   US2003065157-A1.
XX
PD   03-APR-2003.
XX
PF   04-APR-2002; 2002US-00116802.
XX
PR   04-APR-2001; 2001US-0281593P.
XX
PA   (LASE/) LASEK A W.
XX
PI   Lasek AW;
XX
DR   WPI; 2003-540803/51.
XX
PT   New combination comprising cDNAs that are differentially expressed in
PT   respiratory disorders, useful for diagnosing or treating respiratory
PT   disorders e.g., lung cancer, chronic obstructive pulmonary disease,
PT   emphysema or asthma.
XX
PS   Claim 1; Page; 39pp; English.
XX
CC   The invention relates to a combination comprising cDNAs or their
CC   complements that are differentially expressed in respiratory disorder.
CC   The combination is useful for preparing a composition for diagnosing or
CC   treating respiratory disorders e.g. lung cancer, chronic obstructive
CC   pulmonary disease, emphysema or asthma. The present sequence represents
CC   human cDNA differentially expressed during lung cancer
XX
SQ   Sequence 3327 BP; 863 A; 880 C; 834 G; 750 T; 0 U; 0 Other;

  Query Match             79.1%;  Score 17.4;  DB 8;  Length 3327;
  Best Local Similarity   94.7%;  Pred. No. 2.1e+02;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GAGAGTTCTGGGTGTCCTA 22
              |||||||||||| |||||||
Db       1486 GAGAGTTCTGGATGTCCTA 1504


RESULT 12
AAT32455
ID   AAT32455 standard; cDNA; 5149 BP.
XX
AC   AAT32455;
XX
DT   02-DEC-1996  (first entry)
XX
DE   Calpain large subunit 1 coding sequence.
XX
KW   Calpain; subunit; calcium; protease; mutation; treatment; detection;
KW   identification; diagnosis; limb girdle muscular dystrophy; LGMD2;
KW   calcium activated neutral protease; CANP; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    109. .359
FT                   /*tag=  a
FT                   /label= Alu sequence.
FT   CAAT_signal     712. .716
FT                   /*tag=  b
FT                   /note= "Putative CAAT signal."
FT   CAAT_signal     759. .764
FT                   /*tag=  c
FT                   /note= "Putative CAAT signal."
FT   CAAT_signal     979. .983
FT                   /*tag=  d
FT                   /note= "Putative CAAT signal."
FT   TATA_signal     1270. .1275
FT                   /*tag=  e
FT                   /note= "Putative TATA signal."
FT   CDS             1303. .3768
FT                   /*tag=  f
FT                   /product= "Calpain large subunit 1."
FT   polyA_signal    4288. .4293
FT                   /*tag=  g
XX
PN   WO9616175-A2.
XX
PD   30-MAY-1996.
XX
PF   21-NOV-1995;   95WO-EP004575.
XX
PR   22-NOV-1994;   94EP-00402668.
XX
PA   (ASFR-) ASSOC FR CONTRE MYOPATHIES.
XX
PI   Beckmann J,  Richard I;
XX
DR   WPI; 1996-268611/27.
DR   P-PSDB; AAR99579.
XX
PT   Human novel Calpain large sub:unit 1 gene encoding a calcium dependent
PT   protease - used to develop prods. for the diagnosis and treatment of limb
PT   -girdle muscular dystrophy 2 disease.
XX
PS   Claim 1; Fig 2; 66pp; English.
XX
CC   The calpain large subunit 1 gene located on chromosome 15 codes for a
CC   calcium activated neutral protease (CANP3) belonging to the calpain
CC   family. Mutations in the gene induce limb-girdle muscular dystrophy
CC   (LGMD) 2 disease. The gene, and fragments of it, can be used in the
CC   prevention, treatment, diagnosis and detection of a predisposition to
CC   LGMD2 disease
XX
SQ   Sequence 5149 BP; 1343 A; 1312 C; 1179 G; 1314 T; 0 U; 1 Other;

  Query Match             79.1%;  Score 17.4;  DB 2;  Length 5149;
  Best Local Similarity   94.7%;  Pred. No. 2.2e+02;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GAGAGTTCTGGGTGTCCTA 22
              |||||||||||| |||||||
Db       2486 GAGAGTTCTGGATGTCCTA 2504


RESULT 13
AAK67913/c
ID   AAK67913 standard; DNA; 355 BP.
XX
AC   AAK67913;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human immune/haematopoietic antigen genomic sequence SEQ ID NO:22725.
XX
KW   Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW   cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS   Homo sapiens.
XX
PN   WO200157182-A2.
XX
PD   09-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US001354.
```

XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483426/52.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.

XX
PS Disclosure; SEQ ID NO 22725; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX
SQ Sequence 355 BP; 121 A; 77 C; 72 G; 85 T; 0 U; 0 Other;

Query Match 78.2%; Score 17.2; DB 4; Length 355;
Best Local Similarity 86.4%; Pred. No. 2.1e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 TGAGAGAGTTCTGGGTGTCCTA 22
              |||| |||||||| | |||||
Db        146 TGAGTGAGTTCTGGCTCTCCTA 125
```

RESULT 14
AAK67912/c
ID AAK67912 standard; DNA; 356 BP.
XX
AC AAK67912;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:22724.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001354.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.

PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483426/52.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
PS Disclosure; SEQ ID NO 22724; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX

SQ Sequence 356 BP; 121 A; 76 C; 74 G; 85 T; 0 U; 0 Other;

Query Match 78.2%; Score 17.2; DB 4; Length 356;
Best Local Similarity 86.4%; Pred. No. 2.1e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TGAGAGAGTTCTGGGTGTCCTA 22
||||.|||||||||.|.|||||
Db 147 TGAGTGAGTTCTGGCTCTCCTA 126

RESULT 15
AAK56372/c
ID AAK56372 standard; cDNA; 359 BP.
XX
AC AAK56372;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:1432.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ss.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001354.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.

PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483426/52.
DR P-PSDB; AAM83591.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
PS Claim 1; SEQ ID NO 1432; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX
SQ Sequence 359 BP; 122 A; 78 C; 73 G; 84 T; 0 U; 2 Other;

Query Match 78.2%; Score 17.2; DB 4; Length 359;
Best Local Similarity 86.4%; Pred. No. 2.1e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TGAGAGAGTTCTGGGTGTCCTA 22
||||.|||||||||.|.|||||
Db 148 TGAGTGAGTTCTGGCTCTCCTA 127

Search completed: February 19, 2004, 21:51:47
Job time : 24.3821 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 21:21:54 ; Search time 145.664 Seconds
(without alignments)
4510.152 Million cell updates/sec

Title: US-09-939-853A-142
Perfect score: 22
Sequence: 1 tgagagagttctgggtgtccta 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	22	100.0	616	13	BX383606	BX383606 BX383606
	2	22	100.0	778	12	BG178487	BG178487 602328305
	3	22	100.0	878	12	BQ053486	BQ053486 AGENCOURT
	4	22	100.0	986	12	BQ054265	BQ054265 AGENCOURT
	5	22	100.0	1020	12	BQ054281	BQ054281 AGENCOURT
	6	22	100.0	1201	9	AL541041	AL541041 AL541041
	7	18.8	85.5	573	28	CC322697	CC322697 TAM32-33G
	8	18.4	83.6	598	28	AZ827349	AZ827349 2M0103B18
	9	18.4	83.6	651	9	AI727541	AI727541 BNLGHi834
	10	18.4	83.6	2919	11	AK035058	AK035058 Mus muscu
	11	17.8	80.9	283	28	AZ940479	AZ940479 2M0199D23
	12	17.8	80.9	296	29	CC513479	CC513479 CH240_358
	13	17.8	80.9	464	28	BH834762	BH834762 BACPP7-C1
	14	17.8	80.9	498	14	CF606764	CF606764 GEMMA01_0
	15	17.8	80.9	601	28	BH826297	BH826297 BACPP23-O
c	16	17.8	80.9	652	12	BJ011310	BJ011310 BJ011310
c	17	17.8	80.9	690	9	AU168324	AU168324 AU168324
	18	17.8	80.9	699	9	AV846834	AV846834 AV846834
	19	17.8	80.9	780	28	AZ727518	AZ727518 RPCI-24-7
c	20	17.8	80.9	910	14	CA965326	CA965326 CcLL01a04
c	21	17.8	80.9	1062	28	CC305491	CC305491 CH261-59N
c	22	17.8	80.9	1066	28	CC250043	CC250043 CH261-127
	23	17.4	79.1	290	13	BX500783	BX500783 DKFZp779B
c	24	17.4	79.1	295	12	BI430374	BI430374 fq90d03.y
	25	17.4	79.1	438	9	AA128423	AA128423 zn60b10.r
c	26	17.4	79.1	537	9	AV598294	AV598294 AV598294
	27	17.4	79.1	558	28	AZ749071	AZ749071 RPCI-24-7
	28	17.4	79.1	563	29	CE551267	CE551267 tigr-gss-
c	29	17.4	79.1	583	13	BU065505	BU065505 Fgr_7_M01
c	30	17.4	79.1	588	13	BU064867	BU064867 Fgr_6_A04
	31	17.4	79.1	925	14	CF551819	CF551819 AGENCOURT
	32	17.4	79.1	1062	29	CNS06Q2K	AL410130 T3 end of
	33	17.4	79.1	1160	29	CNS06T18	AL413970 T3 end of
	34	17.4	79.1	1878	29	AY419877	AY419877 Homo sapi
	35	17.2	78.2	315	9	AA640882	AA640882 nr24b06.r
	36	17.2	78.2	363	10	BB730751	BB730751 BB730751
c	37	17.2	78.2	383	13	BY675867	BY675867 BY675867
c	38	17.2	78.2	389	29	CG578968	CG578968 OST216823
	39	17.2	78.2	407	9	AV797883	AV797883 AV797883
	40	17.2	78.2	408	10	BE654580	BE654580 UI-M-AJ1-
c	41	17.2	78.2	430	10	BE848016	BE848016 uw31g11.y
	42	17.2	78.2	431	9	AI149606	AI149606 qf31f01.x
	43	17.2	78.2	439	9	AA693188	AA693188 vr62e01.s
	44	17.2	78.2	470	9	AI616154	AI616154 vr62e01.x
	45	17.2	78.2	480	28	AZ264203	AZ264203 RPCI-23-1

ALIGNMENTS

RESULT 1
BX383606
LOCUS BX383606 616 bp mRNA linear EST 08-MAY-2003
DEFINITION BX383606 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
Homo sapiens cDNA clone CS0DJ013YK10 5-PRIME, mRNA sequence.
ACCESSION BX383606
VERSION BX383606.1 GI:30457152
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 616)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 9825.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DJ013BF05QP1&cluster=9825.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ InVitroGen Corporation 1600

```
            Faraday Avenue Genoscope sequence ID : CS0DJ013BF05QP1.
FEATURES             Location/Qualifiers
     source          1. .616
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DJ013YK10"
                     /cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
                     /cell_line="JURKAT"
                     /clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
                     10-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

  Query Match             100.0%;  Score 22;  DB 13;  Length 616;
  Best Local Similarity   100.0%;  Pred. No. 29;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTCCTA 22
              ||||||||||||||||||||||
Db        273 TGAGAGAGTTCTGGGTGTCCTA 294


RESULT 2
BG178487
LOCUS       BG178487                778 bp    mRNA    linear   EST 06-FEB-2001
DEFINITION  602328305F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4429896 5',
            mRNA sequence.
ACCESSION   BG178487
VERSION     BG178487.1  GI:12685190
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 778)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10182  row: i  column: 01
            High quality sequence stop: 657.
FEATURES             Location/Qualifiers
     source          1. .778
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4429896"
                     /tissue_type="adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_91"
                     /note="Organ: prostate; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.4 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             100.0%;  Score 22;  DB 12;  Length 778;
  Best Local Similarity   100.0%;  Pred. No. 31;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTCCTA 22
              ||||||||||||||||||||||
Db        188 TGAGAGAGTTCTGGGTGTCCTA 209


RESULT 3
BQ053486
LOCUS       BQ053486                878 bp    mRNA    linear   EST 29-MAR-2002
DEFINITION  AGENCOURT_6822017 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5935253
            5', mRNA sequence.
ACCESSION   BQ053486
VERSION     BQ053486.1  GI:19812826
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 878)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2122  row: l  column: 06
            High quality sequence stop: 394.
FEATURES             Location/Qualifiers
     source          1. .878
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5935253"
                     /tissue_type="natural killer cells, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_106"
                     /note="Organ: blood; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Library constructed  by Ling Hong in the
                     laboratory of Gerald M. Rubin (University of  California,
                     Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                     Superscript II RT (Life Technologies). Note: this is a
                     NIH_MGC Library."
ORIGIN

  Query Match             100.0%;  Score 22;  DB 12;  Length 878;
  Best Local Similarity   100.0%;  Pred. No. 33;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTCCTA 22
              ||||||||||||||||||||||
Db         99 TGAGAGAGTTCTGGGTGTCCTA 120


RESULT 4
BQ054265
LOCUS       BQ054265                986 bp    mRNA    linear   EST 29-MAR-2002
DEFINITION  AGENCOURT_6830248 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5936339
            5', mRNA sequence.
ACCESSION   BQ054265
VERSION     BQ054265.1  GI:19813605
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 986)
```

```
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2125  row: i  column: 12
            High quality sequence stop: 515.
FEATURES             Location/Qualifiers
     source          1. .986
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5936339"
                     /tissue_type="natural killer cells, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_106"
                     /note="Organ: blood; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Library constructed  by Ling Hong in the
                     laboratory of Gerald M. Rubin (University of  California,
                     Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                     Superscript II RT (Life Technologies). Note: this is a
                     NIH_MGC Library."
ORIGIN

  Query Match           100.0%;  Score 22;  DB 12;  Length 986;
  Best Local Similarity   100.0%;  Pred. No. 34;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTCCTA 22
              ||||||||||||||||||||||
Db        105 TGAGAGAGTTCTGGGTGTCCTA 126


RESULT 5
BQ054281
LOCUS       BQ054281               1020 bp    mRNA    linear   EST 29-MAR-2002
DEFINITION  AGENCOURT_6830234 NIH_MGC_106 Homo sapiens cDNA clone IMAGE:5936362
            5', mRNA sequence.
ACCESSION   BQ054281
VERSION     BQ054281.1  GI:19813621
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1020)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2125  row: j  column: 11
            High quality sequence stop: 556.
FEATURES             Location/Qualifiers
     source          1. .1020
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5936362"
                     /tissue_type="natural killer cells, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_106"
                     /note="Organ: blood; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally cloned
                     into EcoRI/XhoI sites using the following 5' adaptor:
                     GGCACGAG(G). Library constructed  by Ling Hong in the
                     laboratory of Gerald M. Rubin (University of  California,
                     Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
                     Superscript II RT (Life Technologies). Note: this is a
                     NIH_MGC Library."
ORIGIN

  Query Match           100.0%;  Score 22;  DB 12;  Length 1020;
  Best Local Similarity   100.0%;  Pred. No. 34;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTCCTA 22
              ||||||||||||||||||||||
Db         38 TGAGAGAGTTCTGGGTGTCCTA 59


RESULT 6
AL541041
LOCUS       AL541041               1201 bp    mRNA    linear   EST 12-MAY-2003
DEFINITION  AL541041 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE005YK23
            5-PRIME, mRNA sequence.
ACCESSION   AL541041
VERSION     AL541041.2  GI:30544829
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1201)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 15, 2001 this sequence version replaced gi:12871733.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 9825.r For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DE005AF12QP1&cluster=9825.r. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DE005AF12QP1.
FEATURES             Location/Qualifiers
     source          1. .1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DE005YK23"
                     /tissue_type="PLACENTA"
                     /clone_lib="Homo sapiens PLACENTA"
                     /note="Vector: pCMVSPORT_6; 1st strand cDNA was primed
                     with a NotI-oligo(dT) primer. Five prime end enriched,
                     double-strand cDNA was digested with Not I and cloned into
                     the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                     Library was not normalized."
ORIGIN

  Query Match           100.0%;  Score 22;  DB 9;  Length 1201;
  Best Local Similarity   100.0%;  Pred. No. 36;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTCCTA 22
```

```
                    ||||||||||||||||||||
Db        285 TGAGAGAGTTCTGGGTGTCCTA 306


RESULT 7
CC322697
LOCUS       CC322697                573 bp    DNA     linear   GSS 14-MAY-2003
DEFINITION  TAM32-33G24_Sp6.1 TAM32 Gallus gallus genomic clone TAM32-33G24,
            genomic survey sequence.
ACCESSION   CC322697
VERSION     CC322697.1  GI:30716755
KEYWORDS    GSS.
SOURCE      Gallus gallus (chicken)
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 573)
  AUTHORS   Kremitzki,C., Higginbotham,J., Wylie,K., Carter,J., McPherson,J.,
            Warren,W., Graves,T., Mardis,E. and Wilson,R.
  TITLE     Gallus gallus BAC End Reads
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Richard K. Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 152000   Std Error: 0.00
            Seq primer: Sp6 ATTTAGGTGACACTATAG
            Class: BAC ends
            High quality sequence start: 3
            High quality sequence stop: 478.
FEATURES             Location/Qualifiers
     source          1. .573
                     /organism="Gallus gallus"
                     /mol_type="genomic DNA"
                     /strain="Red Jungle Fowl"
                     /db_xref="taxon:9031"
                     /clone="TAM32-33G24"
                     /sex="female"
                     /cell_line="UCD001, inbred 256"
                     /clone_lib="TAM32"
                     /note="Vector: pECBAC1; Site_1: EcoRI; Site_2: EcoRI;
                     TAM32 Female Chicken library - for library and clone
                     ordering information: http://www.hbz.tamu.edu"
ORIGIN

  Query Match             85.5%;  Score 18.8;  DB 28;  Length 573;
  Best Local Similarity   90.9%;  Pred. No. 7.5e+02;
  Matches   20;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTCCTA 22
              |||||||||||||  ||||| ||
Db        361 TGAGAGAGTTCTGTGTGTCTTA 382


RESULT 8
AZ827349
LOCUS       AZ827349                598 bp    DNA     linear   GSS 20-FEB-2001
DEFINITION  2M0103B18R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0103B18 R, genomic survey sequence.
ACCESSION   AZ827349
VERSION     AZ827349.1  GI:12997257
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 598)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0103  row: B  column: 18
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 598.
FEATURES             Location/Qualifiers
     source          1. .598
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0103B18"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
ORIGIN

  Query Match             83.6%;  Score 18.4;  DB 28;  Length 598;
  Best Local Similarity   95.0%;  Pred. No. 1.1e+03;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTCC 20
              |||||||  ||||||||||||
Db        279 TGAGAGAATTCTGGGTGTCC 298


RESULT 9
AI727541
LOCUS       AI727541                651 bp    mRNA    linear   EST 11-JUN-1999
DEFINITION  BNLGHi8347 Six-day Cotton fiber Gossypium hirsutum cDNA 5' similar
            to (AF039373) polyprotein [Arabidopsis thaliana], mRNA sequence.
ACCESSION   AI727541
VERSION     AI727541.1  GI:5046393
KEYWORDS    EST.
SOURCE      Gossypium hirsutum (upland cotton)
  ORGANISM  Gossypium hirsutum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
REFERENCE   1  (bases 1 to 651)
  AUTHORS   Blewitt,M., Matz,E.C., Davy,D.F. and Burr,B.
  TITLE     ESTs from developing cotton fiber
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Ben Burr
```

```
            Biology Department
            Brookhaven National Laboratory
            Upton, NY 11973, USA
            Tel: 516-344-3396
            Fax: 516-344-3407
            Email: burr@bnlux1.bnl.gov
            Seq primer: T3 Primer.
FEATURES              Location/Qualifiers
     source           1. .651
                      /organism="Gossypium hirsutum"
                      /mol_type="mRNA"
                      /cultivar="Acala Maxxa"
                      /db_xref="taxon:3635"
                      /tissue_type="immature fiber"
                      /dev_stage="Six days post anthesis"
                      /lab_host="XL1-Blue"
                      /clone_lib="Six-day Cotton fiber"
                      /note="Vector: pBluescript II KS+"
ORIGIN

  Query Match             83.6%;  Score 18.4;  DB 9;  Length 651;
  Best Local Similarity   95.0%;  Pred. No. 1.2e+03;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTCC 20
              |||||||||||| |||||||
Db         91 TGAGAGAGTTCTTGGTGTCC 110


RESULT 10
AK035058
LOCUS       AK035058               2919 bp    mRNA    linear   HTC 18-SEP-2003
DEFINITION  Mus musculus 12 days embryo embryonic body between diaphragm region
            and neck cDNA, RIKEN full-length enriched library, clone:9430080I20
            product:hypothetical protein, full insert sequence.
ACCESSION   AK035058
VERSION     AK035058.1  GI:26084370
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   The RIKEN Genome Exploration Research Group Phase II Team and the
            FANTOM Consortium.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409, 685-690 (2001)
REFERENCE   5
  AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6  (bases 1 to 2919)
  AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
            Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
            Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
            Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
            Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
            Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
            Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
            Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
            Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
            Please visit our web site for further details.
            URL:http://genome.gsc.riken.go.jp/
            URL:http://fantom.gsc.riken.go.jp/.
FEATURES              Location/Qualifiers
     source           1. .2919
                      /organism="Mus musculus"
                      /mol_type="mRNA"
                      /strain="C57BL/6J"
                      /db_xref="FANTOM_DB:9430080I20"
                      /db_xref="MGI:2399272"
                      /db_xref="taxon:10090"
                      /clone="9430080I20"
                      /tissue_type="embryonic body between diaphragm region and
                      neck"
                      /clone_lib="RIKEN full-length enriched mouse cDNA library"
                      /dev_stage="12 days embryo"
     misc_feature     1. .2919
                      /note="hypothetical protein (evidence: ProCrest)"
ORIGIN

  Query Match             83.6%;  Score 18.4;  DB 11;  Length 2919;
  Best Local Similarity   95.0%;  Pred. No. 1.9e+03;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GAGAGAGTTCTGGGTGTCCT 21
              |||||||||||||||| ||||
Db        589 GAGAGAGTTCTGGGTTTCCT 608


RESULT 11
AZ940479
LOCUS       AZ940479                283 bp    DNA     linear   GSS 26-APR-2001
DEFINITION  2M0199D23R Mouse 10kb plasmid UUGC2M library Mus musculus genomic
            clone UUGC2M0199D23 R, genomic survey sequence.
ACCESSION   AZ940479
VERSION     AZ940479.1  GI:13802090
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 283)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0199  row: D  column: 23
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 283.
FEATURES             Location/Qualifiers
     source          1. .283
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0199D23"
                     /sex="Female"
                     /lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC2M library"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (female) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
ORIGIN

  Query Match            80.9%;  Score 17.8;  DB 28;  Length 283;
  Best Local Similarity  90.5%;  Pred. No. 1.7e+03;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTCCT 21
              || |||||| ||||||||||
Db         50 TGGGAGAGTCCTGGGTGTCCT 70


RESULT 12
CC513479
LOCUS       CC513479                296 bp    DNA     linear   GSS 17-JUN-2003
DEFINITION  CH240_358G5.T7 CHORI-240 Bos taurus genomic clone CH240_358G5,
            genomic survey sequence.
ACCESSION   CC513479
VERSION     CC513479.1  GI:31831767
KEYWORDS    GSS.
SOURCE      Bos taurus (cow)
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 296)
  AUTHORS   Holt,R., Stott,J., Yang,G., Barber,S., Smailus,D., Prabhu,A.-L.,
            Tsai,M., Cloutier,A., Lee,D., Girn,N., Olson,T., Mayo,M.,
            Butterfield,Y., Kirkpatrick,R., Liu,J., Guin,R., Chan,A., Chiu,R.,
            Mathewson,C., Wye,N., Masson,A., Brown-John,M., Jones,S.,
            Schein,J., Marra,M., de Jong,P., McWilliam,S., Barris,W.,
            Dalrymple,B.P. and Tellam,R.
  TITLE     Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398
  JOURNAL   Unpublished (2003)
COMMENT     Other_GSSs: CH240_358G5.TARBAC13P2
            Contact: Rob Holt
            Sequencing
            The British Columbia Cancer Agency Genome Science Centre
            600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6
            Tel: 604-877-6085
            Fax: 604-877-6276
            Email: rholt@bcgsc.ca
            Clones are derived from the bovine BAC library CHORI-240
            (http://www.chori.org/bacpac/bovine240.htm). For BAC library
            availability, please contact Pieter de Jong (pdejong@mail.cho.org).
            Clones may be purchased from BACPAC Resources
            (http://www.chori.org/bacpac/ordering_information.htm). This work
            was undertaken as part of the International Bovine BAC Mapping
            Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
            British Columbia Genome Sciences Centre, Canada.
            Plate: 358  row: G  column: 5
            Seq primer: T7
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .296
                     /organism="Bos taurus"
                     /mol_type="genomic DNA"
                     /strain="breed: Hereford"
                     /db_xref="taxon:9913"
                     /clone="CH240_358G5"
                     /sex="Male"
                     /cell_type="Blood"
                     /clone_lib="CHORI-240"
                     /note="Vector: pTARBAC1.3; Site_1: MboI; Site_2: MboI;
                     Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC
                     library (Male) produced by Pieter de Jong"
ORIGIN

  Query Match            80.9%;  Score 17.8;  DB 29;  Length 296;
  Best Local Similarity  90.5%;  Pred. No. 1.7e+03;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTCCT 21
              | |||||||||||||| |||||
Db         19 TTAGAGAGTTCTGGGAGTCCT 39


RESULT 13
BH834762
LOCUS       BH834762                464 bp    DNA     linear   GSS 20-MAY-2002
DEFINITION  BACPP7-C12.z Pristionchus pacificus BAC ends Pristionchus pacificus
            genomic, genomic survey sequence.
ACCESSION   BH834762
VERSION     BH834762.1  GI:21036033
KEYWORDS    GSS.
SOURCE      Pristionchus pacificus
  ORGANISM  Pristionchus pacificus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
            Neodiplogasteridae; Pristionchus.
REFERENCE   1  (bases 1 to 464)
  AUTHORS   Srinivasan,J., Sinz,W., Lanz,C., Brand,A., Nandakumar,R.,
            Raddatz,G., Witte,H., Keller,H., Kipping,I., Pires da Silva,A.,
            Jesse,T., Millare,J., de Both,M., Schuster,S.C. and Sommer,R.J.
  TITLE     A BAC-based genetic linkage map of the nematode Pristionchus
            pacificus
  JOURNAL   Unpublished (2002)
```

```
COMMENT     Contact: Sommer RJ
            Evolutionary Biology
            Max-Planck-Institute for Developmental Biology
            Spemannstr. 37-39, Tuebingen D-72076, Germany
            Tel: 00497071601371
            Fax: 00497071601498
            Email: ralf.sommer@tuebingen.mpg.de
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .464
                     /organism="Pristionchus pacificus"
                     /mol_type="genomic DNA"
                     /strain="var. California"
                     /db_xref="taxon:54126"
                     /clone_lib="Pristionchus pacificus BAC ends"
ORIGIN

  Query Match             80.9%;  Score 17.8;  DB 28;  Length 464;
  Best Local Similarity   90.5%;  Pred. No. 1.9e+03;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTCCT 21
              ||||||||||| ||| ||||||
Db        315 TGAGAGAGTTATGGCTGTCCT 335


RESULT 14
CF606764
LOCUS       CF606764                498 bp    mRNA    linear   EST 30-SEP-2003
DEFINITION  GEMMA01_000158 Grape Bud pSPORT1 Library Vitis vinifera cDNA 5',
            mRNA sequence.
ACCESSION   CF606764
VERSION     CF606764.1  GI:37187411
KEYWORDS    EST.
SOURCE      Vitis vinifera
  ORGANISM  Vitis vinifera
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; Vitaceae; Vitis.
REFERENCE   1  (bases 1 to 498)
  AUTHORS   Moser,C., Segala,C., Fontana,P., Salakhudtinov,I., Gatto,P.,
            Pindo,M., Zyprian,E., Toepfer,R., Grando,M.S. and Velasco,R.
  TITLE     Expressed sequence tags from different organs of Vitis vinifera
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Moser C
            Laboratorio di Genetica Molecolare
            Istituto Agrario di San Michele all'Adige (IASMA)
            via E. Mach 1, San Michele all'Adige (TN), I-38010, Italia
            Tel: 0039-0461-615314
            Fax: 0039-0461-650956
            Email: claudio.moser@ismaa.it
            The sequencing work has been funded by the 'Fondazione Cassa di
            Risparmio di Trento e Rovereto'
            High quality sequence stop: 498.
FEATURES             Location/Qualifiers
     source          1. .498
                     /organism="Vitis vinifera"
                     /mol_type="mRNA"
                     /cultivar="Pinot noir"
                     /db_xref="taxon:29760"
                     /sex="Hermaphrodite"
                     /dev_stage="bud swelling"
                     /lab_host="DH10B"
                     /clone_lib="Grape Bud pSPORT1 Library"
                     /note="Organ: bud; Vector: pSPORT1; Site_1: NotI; Site_2:
                     SalI"
ORIGIN

  Query Match             80.9%;  Score 17.8;  DB 14;  Length 498;
  Best Local Similarity   90.5%;  Pred. No. 2e+03;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GAGAGAGTTCTGGGTGTCCTA 22
              ||| ||| |||||||||||||
Db         88 GAGTGAGATCTGGGTGTCCTA 108


RESULT 15
BH826297
LOCUS       BH826297                601 bp    DNA     linear   GSS 20-MAY-2002
DEFINITION  BACPP23-O08.z Pristionchus pacificus BAC ends Pristionchus
            pacificus genomic, genomic survey sequence.
ACCESSION   BH826297
VERSION     BH826297.1  GI:21006331
KEYWORDS    GSS.
SOURCE      Pristionchus pacificus
  ORGANISM  Pristionchus pacificus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
            Neodiplogasteridae; Pristionchus.
REFERENCE   1  (bases 1 to 601)
  AUTHORS   Srinivasan,J., Sinz,W., Lanz,C., Brand,A., Nandakumar,R.,
            Raddatz,G., Witte,H., Keller,H., Kipping,I., Pires da Silva,A.,
            Jesse,T., Millare,J., de Both,M., Schuster,S.C. and Sommer,R.J.
  TITLE     A BAC-based genetic linkage map of the nematode Pristionchus
            pacificus
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Sommer RJ
            Evolutionary Biology
            Max-Planck-Institute for Developmental Biology
            Spemannstr. 37-39, Tuebingen D-72076, Germany
            Tel: 00497071601371
            Fax: 00497071601498
            Email: ralf.sommer@tuebingen.mpg.de
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .601
                     /organism="Pristionchus pacificus"
                     /mol_type="genomic DNA"
                     /strain="var. California"
                     /db_xref="taxon:54126"
                     /clone_lib="Pristionchus pacificus BAC ends"
ORIGIN

  Query Match             80.9%;  Score 17.8;  DB 28;  Length 601;
  Best Local Similarity   90.5%;  Pred. No. 2.1e+03;
  Matches   19;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTCCT 21
              ||||||||||| ||| ||||||
Db        315 TGAGAGAGTTATGGCTGTCCT 335


Search completed: February 20, 2004, 01:41:44
Job time : 149.664 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 21:32:00 ; Search time 3.56994 Seconds
(without alignments)
3419.919 Million cell updates/sec

Title: US-09-939-853A-142
Perfect score: 22
Sequence: 1 tgagagagttctgggtgtccta 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	17.2	78.2	1245	4	US-09-489-039A-3698	Sequence 3698, Ap
	2	16.4	74.5	1273	3	US-08-725-758A-3	Sequence 3, Appli
	3	16.4	74.5	1373	3	US-08-725-758A-1	Sequence 1, Appli
c	4	16.2	73.6	618	4	US-09-621-976-788	Sequence 788, App
	5	16.2	73.6	2713	2	US-08-916-901-6	Sequence 6, Appli
	6	16.2	73.6	2713	4	US-09-154-602-6	Sequence 6, Appli
c	7	15.8	71.8	274	4	US-09-313-294A-5461	Sequence 5461, Ap
c	8	15.8	71.8	283	4	US-09-313-294A-4815	Sequence 4815, Ap
c	9	15.8	71.8	288	4	US-09-313-294A-809	Sequence 809, App
c	10	15.8	71.8	288	4	US-09-313-294A-2911	Sequence 2911, Ap
	11	15.8	71.8	835	4	US-09-833-381-1328	Sequence 1328, Ap
	12	15.8	71.8	2061	4	US-09-653-839-7	Sequence 7, Appli
	13	15.8	71.8	2109	4	US-09-653-839-5	Sequence 5, Appli
	14	15.8	71.8	2172	4	US-09-653-839-3	Sequence 3, Appli
	15	15.8	71.8	2220	4	US-09-653-839-1	Sequence 1, Appli
	16	15.8	71.8	2353	4	US-09-622-880B-2	Sequence 2, Appli
	17	15.8	71.8	2806	4	US-09-653-839-9	Sequence 9, Appli
	18	15.8	71.8	3138	4	US-09-622-880B-16	Sequence 16, Appl
c	19	15.6	70.9	291	4	US-09-313-294A-6747	Sequence 6747, Ap
	20	15.6	70.9	331	4	US-09-621-976-11814	Sequence 11814, A
	21	15.6	70.9	331	4	US-09-621-976-11984	Sequence 11984, A
	22	15.6	70.9	344	4	US-09-621-976-12361	Sequence 12361, A
	23	15.6	70.9	344	4	US-09-621-976-13537	Sequence 13537, A
	24	15.6	70.9	7400	1	US-07-674-852-1	Sequence 1, Appli
	25	15.6	70.9	7400	3	US-08-473-185-1	Sequence 1, Appli
	26	15.6	70.9	7400	3	US-09-171-387-3	Sequence 3, Appli
	27	15.6	70.9	9573	4	US-09-220-132-168	Sequence 168, App
	28	15.4	70.0	2972	2	US-08-720-484A-3	Sequence 3, Appli
	29	15.4	70.0	2972	3	US-08-953-823A-3	Sequence 3, Appli
	30	15.4	70.0	2972	4	US-09-398-239-3	Sequence 3, Appli
	31	15.4	70.0	2972	4	US-09-560-876A-3	Sequence 3, Appli
c	32	15.4	70.0	3301	4	US-09-148-545-66	Sequence 66, Appl
	33	15.2	69.1	747	4	US-08-630-915A-39	Sequence 39, Appl
c	34	15.2	69.1	1167	4	US-09-679-686B-9	Sequence 9, Appli
c	35	15.2	69.1	1472	1	US-08-123-161A-9	Sequence 9, Appli
c	36	15.2	69.1	1472	1	US-08-483-278-9	Sequence 9, Appli
c	37	15.2	69.1	1558	4	US-09-123-030-7	Sequence 7, Appli
	38	15.2	69.1	2389	3	US-08-691-563C-52	Sequence 52, Appl
	39	15.2	69.1	2389	4	US-09-374-766-52	Sequence 52, Appl
	40	15.2	69.1	2389	4	US-08-979-847B-48	Sequence 48, Appl
c	41	15.2	69.1	2464	4	US-09-620-312D-448	Sequence 448, App
c	42	15.2	69.1	2584	3	US-08-758-662-8	Sequence 8, Appli
	43	15.2	69.1	2873	4	US-08-630-915A-193	Sequence 193, App
c	44	15.2	69.1	2905	4	US-09-595-684B-24	Sequence 24, Appl
c	45	15.2	69.1	3051	4	US-09-409-604-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-489-039A-3698
; Sequence 3698, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 3698
; LENGTH: 1245
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-3698

Query Match 78.2%; Score 17.2; DB 4; Length 1245;
Best Local Similarity 86.4%; Pred. No. 30;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 TGAGAGAGTTCTGGGTGTCCTA 22
              |||| || |||||| |||||||
Db        800 TGAGCGATTTCTGGATGTCCTA 821
```

RESULT 2
US-08-725-758A-3
; Sequence 3, Application US/08725758A
; Patent No. 6160108
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy
; APPLICANT: Clement, Christophe Y.
; TITLE OF INVENTION: NOVEL PLATELET ACTIVATION PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/725,758A
;      FILING DATE:  04-OCT-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/005,074
;      FILING DATE:  06-OCT-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fraser, Janis K.
;      REGISTRATION NUMBER:  34,819
;      REFERENCE/DOCKET NUMBER:  05433/020001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-542-5070
;      TELEFAX:  617-542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1273 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  Coding Sequence
;      LOCATION:  44...1273
US-08-725-758A-3

  Query Match             74.5%;  Score 16.4;  DB 3;  Length 1273;
  Best Local Similarity   94.4%;  Pred. No. 71;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGT 18
              ||||||||||| ||||||
Db        372 TGAGAGAGTTCCGGGTGT 389


RESULT 3
US-08-725-758A-1
; Sequence 1, Application US/08725758A
; Patent No. 6160108
;  GENERAL INFORMATION:
;    APPLICANT:  Reed, Guy
;    APPLICANT:  Clement, Christophe Y.
;    TITLE OF INVENTION:  NOVEL PLATELET ACTIVATION PROTEIN
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/725,758A
;      FILING DATE:  04-OCT-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/005,074
;      FILING DATE:  06-OCT-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fraser, Janis K.
;      REGISTRATION NUMBER:  34,819
;      REFERENCE/DOCKET NUMBER:  05433/020001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-542-5070
;      TELEFAX:  617-542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1373 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  Coding Sequence
;      LOCATION:  44...1321
US-08-725-758A-1

  Query Match             74.5%;  Score 16.4;  DB 3;  Length 1373;
  Best Local Similarity   94.4%;  Pred. No. 72;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGT 18
              ||||||||||| ||||||
Db        372 TGAGAGAGTTCCGGGTGT 389


RESULT 4
US-09-621-976-788/c
; Sequence 788, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
;  APPLICANT: Dumas Milne Edwards, J.B.
;  APPLICANT:  Jobert, S.
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: ESTs and Encoded Human Proteins.
;  FILE REFERENCE: GENSET.054PR2
;  CURRENT APPLICATION NUMBER: US/09/621,976
;  CURRENT FILING DATE:  2000-07-21
;  NUMBER OF SEQ ID NOS: 19335
;  SOFTWARE: Patent.pm
; SEQ ID NO 788
;   LENGTH: 618
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 306..617
US-09-621-976-788

  Query Match             73.6%;  Score 16.2;  DB 4;  Length 618;
  Best Local Similarity   85.7%;  Pred. No. 83;
  Matches   18;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTCCT 21
              | |||| ||||| ||||||||
Db        540 TCAGAGGGTTCTCGGTGTCCT 520


RESULT 5
US-08-916-901-6
; Sequence 6, Application US/08916901
; Patent No. 5892012
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Corley, Neil C.
;    APPLICANT:  Shah, Purvi
;    TITLE OF INVENTION:  RAB PROTEINS
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Dr.
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
```

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/916,901
;      FILING DATE:  Filed Herewith
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0367 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2713 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-916-901-6

  Query Match             73.6%;  Score 16.2;  DB 2;  Length 2713;
  Best Local Similarity   85.7%;  Pred. No. 96;
  Matches   18;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTCCT 21
              |||| ||||| ||| ||||||
Db       2401 TGAGTGAGTTTTGGATGTCCT 2421


RESULT 6
US-09-154-602-6
; Sequence 6, Application US/09154602
; Patent No. 6300472
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Corley, Neil C.
;    APPLICANT:  Shah, Purvi
;    TITLE OF INVENTION:  RAB PROTEINS
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Dr.
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/154,602
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/916,901
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0367 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2713 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-154-602-6

  Query Match             73.6%;  Score 16.2;  DB 4;  Length 2713;
  Best Local Similarity   85.7%;  Pred. No. 96;
  Matches   18;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTCCT 21
              |||| ||||| ||| ||||||
Db       2401 TGAGTGAGTTTTGGATGTCCT 2421


RESULT 7
US-09-313-294A-5461/c
; Sequence 5461, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
;  APPLICANT:  Lalgudi, Raghunath V.
;  APPLICANT:  Ito, Laura Y.
;  APPLICANT:  Sherman, Bradley K.
;  TITLE OF INVENTION:  POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
;  FILE REFERENCE:  PL-0017 US
;  CURRENT APPLICATION NUMBER: US/09/313,294A
;  CURRENT FILING DATE:  1999-05-14
;  NUMBER OF SEQ ID NOS:  7600
;  SOFTWARE:  PERL Program
; SEQ ID NO 5461
;   LENGTH: 274
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6476212 700350185H1
US-09-313-294A-5461

  Query Match             71.8%;  Score 15.8;  DB 4;  Length 274;
  Best Local Similarity   89.5%;  Pred. No. 1.2e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTC 19
              ||| |||||||||||||  |
Db         36 TGATAGAGTTCTGGGTGGC 18


RESULT 8
US-09-313-294A-4815/c
; Sequence 4815, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
;  APPLICANT:  Lalgudi, Raghunath V.
;  APPLICANT:  Ito, Laura Y.
;  APPLICANT:  Sherman, Bradley K.
;  TITLE OF INVENTION:  POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
;  FILE REFERENCE:  PL-0017 US
;  CURRENT APPLICATION NUMBER: US/09/313,294A
;  CURRENT FILING DATE:  1999-05-14
;  NUMBER OF SEQ ID NOS:  7600
;  SOFTWARE:  PERL Program
; SEQ ID NO 4815
;   LENGTH: 283
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6476212 700349077H1
;   NAME/KEY: unsure
;   LOCATION: 14, 132
;   OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-4815

  Query Match             71.8%;  Score 15.8;  DB 4;  Length 283;
  Best Local Similarity   89.5%;  Pred. No. 1.2e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTC 19
```

```
              ||| |||||||||||||| |
Db        259 TGATAGAGTTCTGGGTGGC 241


RESULT 9
US-09-313-294A-809/c
; Sequence 809, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
;  APPLICANT:  Lalgudi, Raghunath V.
;  APPLICANT:  Ito, Laura Y.
;  APPLICANT:  Sherman, Bradley K.
;  TITLE OF INVENTION:  POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
;  FILE REFERENCE:  PL-0017 US
;  CURRENT APPLICATION NUMBER: US/09/313,294A
;  CURRENT FILING DATE:  1999-05-14
;  NUMBER OF SEQ ID NOS:  7600
;  SOFTWARE:  PERL Program
; SEQ ID NO 809
;   LENGTH: 288
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6476212 700549871H1
US-09-313-294A-809

  Query Match             71.8%;  Score 15.8;  DB 4;  Length 288;
  Best Local Similarity   89.5%;  Pred. No. 1.2e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTC 19
              ||| |||||||||||||| |
Db        287 TGATAGAGTTCTGGGTGGC 269


RESULT 10
US-09-313-294A-2911/c
; Sequence 2911, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
;  APPLICANT:  Lalgudi, Raghunath V.
;  APPLICANT:  Ito, Laura Y.
;  APPLICANT:  Sherman, Bradley K.
;  TITLE OF INVENTION:  POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
;  FILE REFERENCE:  PL-0017 US
;  CURRENT APPLICATION NUMBER: US/09/313,294A
;  CURRENT FILING DATE:  1999-05-14
;  NUMBER OF SEQ ID NOS:  7600
;  SOFTWARE:  PERL Program
; SEQ ID NO 2911
;   LENGTH: 288
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. 6476212 700553476H1
US-09-313-294A-2911

  Query Match             71.8%;  Score 15.8;  DB 4;  Length 288;
  Best Local Similarity   89.5%;  Pred. No. 1.2e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTC 19
              ||| |||||||||||||| |
Db        260 TGATAGAGTTCTGGGTGGC 242


RESULT 11
US-09-833-381-1328
; Sequence 1328, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
;  APPLICANT: Robison, Keith E.
;  TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
;  FILE REFERENCE: 5800-119
;  CURRENT APPLICATION NUMBER: US/09/833,381
;  CURRENT FILING DATE:  2001-04-11
;  PRIOR APPLICATION NUMBER: 09/516,448
;  PRIOR FILING DATE: 2000-02-29
;  NUMBER OF SEQ ID NOS: 2050
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1328
;   LENGTH: 835
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-833-381-1328

  Query Match             71.8%;  Score 15.8;  DB 4;  Length 835;
  Best Local Similarity   89.5%;  Pred. No. 1.3e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 AGAGAGTTCTGGGTGTCCT 21
              ||| |||||||| ||||||
Db        755 AGACAGTTCTGGTTGTCCT 773


RESULT 12
US-09-653-839-7
; Sequence 7, Application US/09653839
; Patent No. 6433153
; GENERAL INFORMATION:
;  APPLICANT: Donoho, Gregory
;  APPLICANT:  Turner, C. Alexander Jr.
;  APPLICANT:  Nehls, Michael
;  APPLICANT:  Friedrich, Glenn
;  APPLICANT:  Zambrowicz, Brian
;  APPLICANT:  Sands, Arthur T.
;  TITLE OF INVENTION: No. 6433153el Human Calcium Dependent Proteases
;  TITLE OF INVENTION:  and Polynucleotides Encoding the Same
;  FILE REFERENCE: LEX-0038-USA
;  CURRENT APPLICATION NUMBER: US/09/653,839
;  CURRENT FILING DATE:  2000-09-01
;  PRIOR APPLICATION NUMBER: US 60/152,057
;  PRIOR FILING DATE: 1999-09-02
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
;   LENGTH: 2061
;   TYPE: DNA
;   ORGANISM: homo sapiens
US-09-653-839-7

  Query Match             71.8%;  Score 15.8;  DB 4;  Length 2061;
  Best Local Similarity   89.5%;  Pred. No. 1.4e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 GAGAGTTCTGGGTGTCCTA 22
              | |||||||||| |||||||
Db        956 GGGAGTTCTGGATGTCCTA 974


RESULT 13
US-09-653-839-5
; Sequence 5, Application US/09653839
; Patent No. 6433153
; GENERAL INFORMATION:
;  APPLICANT: Donoho, Gregory
;  APPLICANT:  Turner, C. Alexander Jr.
;  APPLICANT:  Nehls, Michael
;  APPLICANT:  Friedrich, Glenn
;  APPLICANT:  Zambrowicz, Brian
;  APPLICANT:  Sands, Arthur T.
;  TITLE OF INVENTION: No. 6433153el Human Calcium Dependent Proteases
;  TITLE OF INVENTION:  and Polynucleotides Encoding the Same
;  FILE REFERENCE: LEX-0038-USA
```

```
;  CURRENT APPLICATION NUMBER: US/09/653,839
;  CURRENT FILING DATE:  2000-09-01
;  PRIOR APPLICATION NUMBER: US 60/152,057
;  PRIOR FILING DATE: 1999-09-02
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 2109
;   TYPE: DNA
;   ORGANISM: homo sapiens
US-09-653-839-5

  Query Match             71.8%;  Score 15.8;  DB 4;  Length 2109;
  Best Local Similarity   89.5%;  Pred. No. 1.4e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 GAGAGTTCTGGGTGTCCTA 22
              | ||||||||| |||||||
Db        956 GGGAGTTCTGGATGTCCTA 974


RESULT 14
US-09-653-839-3
; Sequence 3, Application US/09653839
; Patent No. 6433153
; GENERAL INFORMATION:
;  APPLICANT: Donoho, Gregory
;  APPLICANT:  Turner, C. Alexander Jr.
;  APPLICANT:  Nehls, Michael
;  APPLICANT:  Friedrich, Glenn
;  APPLICANT:  Zambrowicz, Brian
;  APPLICANT:  Sands, Arthur T.
;  TITLE OF INVENTION: No. 6433153el Human Calcium Dependent Proteases
;  TITLE OF INVENTION:  and Polynucleotides Encoding the Same
;  FILE REFERENCE: LEX-0038-USA
;  CURRENT APPLICATION NUMBER: US/09/653,839
;  CURRENT FILING DATE:  2000-09-01
;  PRIOR APPLICATION NUMBER: US 60/152,057
;  PRIOR FILING DATE: 1999-09-02
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 2172
;   TYPE: DNA
;   ORGANISM: homo sapiens
US-09-653-839-3

  Query Match             71.8%;  Score 15.8;  DB 4;  Length 2172;
  Best Local Similarity   89.5%;  Pred. No. 1.4e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 GAGAGTTCTGGGTGTCCTA 22
              | ||||||||| |||||||
Db       1067 GGGAGTTCTGGATGTCCTA 1085


RESULT 15
US-09-653-839-1
; Sequence 1, Application US/09653839
; Patent No. 6433153
; GENERAL INFORMATION:
;  APPLICANT: Donoho, Gregory
;  APPLICANT:  Turner, C. Alexander Jr.
;  APPLICANT:  Nehls, Michael
;  APPLICANT:  Friedrich, Glenn
;  APPLICANT:  Zambrowicz, Brian
;  APPLICANT:  Sands, Arthur T.
;  TITLE OF INVENTION: No. 6433153el Human Calcium Dependent Proteases
;  TITLE OF INVENTION:  and Polynucleotides Encoding the Same
;  FILE REFERENCE: LEX-0038-USA
;  CURRENT APPLICATION NUMBER: US/09/653,839
;  CURRENT FILING DATE:  2000-09-01
;  PRIOR APPLICATION NUMBER: US 60/152,057
;  PRIOR FILING DATE: 1999-09-02
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 2220
;   TYPE: DNA
;   ORGANISM: homo sapiens
US-09-653-839-1

  Query Match             71.8%;  Score 15.8;  DB 4;  Length 2220;
  Best Local Similarity   89.5%;  Pred. No. 1.5e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 GAGAGTTCTGGGTGTCCTA 22
              | ||||||||| |||||||
Db       1067 GGGAGTTCTGGATGTCCTA 1085


Search completed: February 20, 2004, 01:45:12
Job time : 4.56994 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: February 19, 2004, 23:23:34 ; Search time 52.8809 Seconds
(without alignments)
1456.787 Million cell updates/sec

Title: US-09-939-853A-142
Perfect score: 22
Sequence: 1 tgagagagttctgggtgtccta 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2308684 seqs, 1750822206 residues

Total number of hits satisfying chosen parameters: 4617368

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	22	100.0	864	10	US-09-814-353-21302	Sequence 21302, A
	2	20.4	92.7	763	9	US-09-867-550-953	Sequence 953, App
c	3	17.8	80.9	2064	15	US-10-108-260A-1362	Sequence 1362, Ap
	4	17.4	79.1	2442	15	US-10-080-334-85	Sequence 85, Appl
	5	17.4	79.1	2466	15	US-10-159-563-396	Sequence 396, App
	6	17.4	79.1	2789	15	US-10-274-639-22	Sequence 22, Appl
	7	17.2	78.2	545	15	US-10-027-632-284738	Sequence 284738,
c	8	17.2	78.2	641	15	US-10-027-632-131511	Sequence 131511,
	9	17.2	78.2	904	15	US-10-027-632-131512	Sequence 131512,
	10	17.2	78.2	1163	12	US-10-282-122A-37180	Sequence 37180, A
	11	17.2	78.2	1182	12	US-10-282-122A-38988	Sequence 38988, A
	12	17.2	78.2	1182	12	US-10-282-122A-39447	Sequence 39447, A
	13	17.2	78.2	1185	12	US-10-282-122A-23260	Sequence 23260, A
c	14	16.8	76.4	2394	15	US-10-104-047-1043	Sequence 1043, Ap
c	15	16.8	76.4	301692	15	US-10-428-487-11	Sequence 11, Appl
	16	16.4	74.5	1191	15	US-10-264-237-915	Sequence 915, App
	17	16.4	74.5	1357	10	US-09-890-688-85	Sequence 85, Appl
	18	16.4	74.5	1621	9	US-09-729-674-171	Sequence 171, App
c	19	16.4	74.5	2200	15	US-10-108-260A-1744	Sequence 1744, Ap
	20	16.4	74.5	2560	12	US-10-424-599-130837	Sequence 130837,
c	21	16.4	74.5	3128	10	US-09-919-039-373	Sequence 373, App
c	22	16.4	74.5	3128	14	US-10-101-510-583	Sequence 583, App
	23	16.4	74.5	27684	15	US-10-034-650-28	Sequence 28, Appl
	24	16.2	73.6	179	14	US-10-029-386-17034	Sequence 17034, A
	25	16.2	73.6	279	14	US-10-029-386-19880	Sequence 19880, A
	26	16.2	73.6	294	9	US-09-864-761-19936	Sequence 19936, A
	27	16.2	73.6	297	9	US-09-796-692-9356	Sequence 9356, Ap
	28	16.2	73.6	297	14	US-10-040-862-9356	Sequence 9356, Ap
	29	16.2	73.6	297	15	US-10-057-475B-9356	Sequence 9356, Ap
	30	16.2	73.6	297	15	US-10-154-884B-9356	Sequence 9356, Ap
	31	16.2	73.6	444	9	US-09-864-761-3157	Sequence 3157, Ap
	32	16.2	73.6	457	15	US-10-027-632-56117	Sequence 56117, A
	33	16.2	73.6	457	15	US-10-027-632-299908	Sequence 299908,
	34	16.2	73.6	502	15	US-10-027-632-61326	Sequence 61326, A
	35	16.2	73.6	525	14	US-10-029-386-20660	Sequence 20660, A
	36	16.2	73.6	542	10	US-09-991-936-1497	Sequence 1497, Ap
	37	16.2	73.6	542	14	US-10-029-386-6149	Sequence 6149, Ap
	38	16.2	73.6	555	14	US-10-029-386-22402	Sequence 22402, A
c	39	16.2	73.6	555	14	US-10-029-386-24654	Sequence 24654, A
	40	16.2	73.6	593	14	US-10-029-386-3334	Sequence 3334, Ap
	41	16.2	73.6	596	14	US-10-029-386-8674	Sequence 8674, Ap
c	42	16.2	73.6	596	14	US-10-029-386-10951	Sequence 10951, A
c	43	16.2	73.6	620	15	US-10-027-632-213371	Sequence 213371,
c	44	16.2	73.6	620	15	US-10-027-632-213372	Sequence 213372,
c	45	16.2	73.6	647	15	US-10-027-632-275671	Sequence 275671,

ALIGNMENTS

RESULT 1
US-09-814-353-21302
; Sequence 21302, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 21302
; LENGTH: 864
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 2, 3, 32, 862, 863, 864
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-21302

Query Match 100.0%; Score 22; DB 10; Length 864;
Best Local Similarity 100.0%; Pred. No. 0.34;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TGAGAGAGTTCTGGGTGTCCTA 22
              ||||||||||||||||||||||
Db        276 TGAGAGAGTTCTGGGTGTCCTA 297
```

RESULT 2
US-09-867-550-953
; Sequence 953, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells and
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 953
; LENGTH: 763
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-867-550-953

Query Match 92.7%; Score 20.4; DB 9; Length 763;
Best Local Similarity 95.5%; Pred. No. 2.1;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 TGAGAGAGTTCTGGGTGTCCTA 22
              |||||||||||||| |||||||
Db        112 TGAGAGAGTTCTGGATGTCCTA 133
```

RESULT 3
US-10-108-260A-1362/c
; Sequence 1362, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1362
; LENGTH: 2064
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-1362

Query Match 80.9%; Score 17.8; DB 15; Length 2064;
Best Local Similarity 90.5%; Pred. No. 41;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 TGAGAGAGTTCTGGGTGTCCT 21
              ||||||||| |||||||| ||
Db        159 TGAGAGAGTGCTGGGTGTTCT 139
```

RESULT 4
US-10-080-334-85
; Sequence 85, Application US/10080334
; Publication No. US20040002584A1
; GENERAL INFORMATION:
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shimkets, Richard A
; APPLICANT: Li, Li
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Vernet, Corine A. M.
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Guo, Xiaojia
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc L
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Patturajan, Meera
; APPLICANT: Gangolli, Esha A
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Liu, Xiaohong
; APPLICANT: Baumgartner, Jason C.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Spaderna, Steven K
; APPLICANT: Zerhusen, Bryan D
; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
; TITLE OF INVENTION: Using the Same
; FILE REFERENCE: 21402-275
; CURRENT APPLICATION NUMBER: US/10/080,334
; CURRENT FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/270,523
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/322,712
; PRIOR FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: 60/311,980
; PRIOR FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: 60/330,307
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/278,796
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: 60/281,521
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/276,677
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/311,595
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/270,220
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/274,295
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/318,526
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/286,548
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/291,765
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 60/270,797
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 60/276,400
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/270,810
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 85
; LENGTH: 2442
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-080-334-85

Query Match 79.1%; Score 17.4; DB 15; Length 2442;
Best Local Similarity 94.7%; Pred. No. 64;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          4 GAGAGTTCTGGGTGTCCTA 22
              ||||||||||| |||||||
Db       1116 GAGAGTTCTGGATGTCCTA 1134


RESULT 5
US-10-159-563-396
; Sequence 396, Application US/10159563
; Publication No. US20040009154A1
; GENERAL INFORMATION:
;  APPLICANT: Khan, Javed
;  APPLICANT:  Ringner, Markus
;  APPLICANT:  Peterson, Carsten
;  APPLICANT:  Meltzer, Paul
;  TITLE OF INVENTION: SELECTIONS OF GENES AND METHODS OF USING THE SAME FOR
;  TITLE OF INVENTION:  DIAGNOSIS AND FOR TARGETING THE THERAPY OF SELECT CANCERS
;  FILE REFERENCE: 11613.56USI1
;  CURRENT APPLICATION NUMBER: US/10/159,563
;  CURRENT FILING DATE:  2002-12-09
;  PRIOR APPLICATION NUMBER: US 10/133,937
;  PRIOR FILING DATE: 2002-04-25
;  NUMBER OF SEQ ID NOS: 444
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 396
;   LENGTH: 2466
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-159-563-396

  Query Match             79.1%;  Score 17.4;  DB 15;  Length 2466;
  Best Local Similarity   94.7%;  Pred. No. 64;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GAGAGTTCTGGGTGTCCTA 22
              ||||||||||| |||||||
Db       1184 GAGAGTTCTGGATGTCCTA 1202


RESULT 6
US-10-274-639-22
; Sequence 22, Application US/10274639
; Publication No. US20030232349A1
; GENERAL INFORMATION:
;  APPLICANT: INCYTE GENOMICS, INC.
;  APPLICANT:  DELEGEANE, Angelo M.; GANDHI, Ameena R.
;  APPLICANT:  HAFALIA, April J.A.; LU, Dyung Aina M.
;  APPLICANT:  PATTERSON, Chandra; TRIBOULEY, Catherine M.
;  APPLICANT:  DAS, Debopriya; KALLICK, Deborah A.
;  APPLICANT:  NGUYEN, Danniel B.; LEE, Ernestine A.
;  APPLICANT:  KHAN, Farrah A.; YUE, Henry
;  APPLICANT:  AU-YOUNG, Janice K.; GRIFFIN, Jennifer A.
;  APPLICANT:  POLICKY, Jennifer L.; RAMKUMAR, Jayalaxmi
;  APPLICANT:  YANG, Junming; THANGAVELU, Kavitha
;  APPLICANT:  DING, Li; KEARNEY, Liam
;  APPLICANT:  BAUGHN, Mariah R.; BOROWSKY, Mark L.
;  APPLICANT:  SANJANWALA, Madhusudan M.; YAO, Monique G.
;  APPLICANT:  BURFORD, Neil; WALIA, Narinder K.
;  APPLICANT:  LAL, Preeti G.; LEE, Sally
;  APPLICANT:  TODD, Stephen; LO, Terence P.
;  APPLICANT:  TANG, Y. Tom; ELLIOTT, Vicki S.
;  APPLICANT:  AZIMZAI, Yalda; LU, Yan
;  TITLE OF INVENTION: PROTEASES
;  FILE REFERENCE: PI-0167 USA
;  CURRENT APPLICATION NUMBER: US/10/274,639
;  CURRENT FILING DATE:  2002-10-18
;  PRIOR APPLICATION NUMBER: PCT/US01/22397
;  PRIOR FILING DATE: 2001-07-17
;  PRIOR APPLICATION NUMBER: US 60/220,063
;  PRIOR FILING DATE: 2000-07-21
;  PRIOR APPLICATION NUMBER: US 60/221,680
;  PRIOR FILING DATE: 2000-07-28
;  PRIOR APPLICATION NUMBER: US 60/223,544
;  PRIOR FILING DATE: 2000-08-04
;  PRIOR APPLICATION NUMBER: US 60/224,717
;  PRIOR FILING DATE: 2000-08-11
;  PRIOR APPLICATION NUMBER: US 60/225,988
;  PRIOR FILING DATE: 2000-08-16
;  PRIOR APPLICATION NUMBER: US 60/227,568
;  PRIOR FILING DATE: 2000-08-23
;  NUMBER OF SEQ ID NOS: 42
;  SOFTWARE: PERL Program
; SEQ ID NO 22
;   LENGTH: 2789
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20030232349A1 5155802CB1
US-10-274-639-22

  Query Match             79.1%;  Score 17.4;  DB 15;  Length 2789;
  Best Local Similarity   94.7%;  Pred. No. 64;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GAGAGTTCTGGGTGTCCTA 22
              ||||||||||| |||||||
Db       1342 GAGAGTTCTGGATGTCCTA 1360


RESULT 7
US-10-027-632-284738
; Sequence 284738, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 284738
;   LENGTH: 545
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-284738

  Query Match             78.2%;  Score 17.2;  DB 15;  Length 545;
  Best Local Similarity   86.4%;  Pred. No. 79;
  Matches   19;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTCCTA 22
              ||||| | |||||||||||| ||
Db        196 TGAGACATTTCTGGGTGTCATA 217


RESULT 8
US-10-027-632-131511/c
; Sequence 131511, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
```

```
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 131511
;   LENGTH: 641
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-131511

  Query Match             78.2%;  Score 17.2;  DB 15;  Length 641;
  Best Local Similarity   86.4%;  Pred. No. 80;
  Matches   19;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTCCTA 22
              ||||||| ||||||||| || |
Db        371 TGAGAGACTTCTGGGTGACCCA 350


RESULT 9
US-10-027-632-131512
; Sequence 131512, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 131512
;   LENGTH: 904
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-131512

  Query Match             78.2%;  Score 17.2;  DB 15;  Length 904;
  Best Local Similarity   86.4%;  Pred. No. 80;
  Matches   19;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTCCTA 22
              ||||||| ||||||||| || |
Db        272 TGAGAGACTTCTGGGTGACCCA 293


RESULT 10
US-10-282-122A-37180
; Sequence 37180, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 37180
;   LENGTH: 1163
;   TYPE: DNA
;   ORGANISM: Salmonella paratyphi A
US-10-282-122A-37180

  Query Match             78.2%;  Score 17.2;  DB 12;  Length 1163;
  Best Local Similarity   86.4%;  Pred. No. 80;
  Matches   19;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTCCTA 22
              |||| || |||||| |||||||
Db        739 TGAGTGATTTCTGGATGTCCTA 760


RESULT 11
US-10-282-122A-38988
; Sequence 38988, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
```

```
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 38988
;   LENGTH: 1182
;   TYPE: DNA
;   ORGANISM: Salmonella typhimurium
US-10-282-122A-38988

  Query Match             78.2%;  Score 17.2;  DB 12;  Length 1182;
  Best Local Similarity   86.4%;  Pred. No. 80;
  Matches   19;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTCCTA 22
              |||| || |||||| |||||||
Db        740 TGAGTGATTTCTGGATGTCCTA 761


RESULT 12
US-10-282-122A-39447
; Sequence 39447, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 39447
;   LENGTH: 1182
;   TYPE: DNA
;   ORGANISM: Salmonella typhi
US-10-282-122A-39447

  Query Match             78.2%;  Score 17.2;  DB 12;  Length 1182;
  Best Local Similarity   86.4%;  Pred. No. 80;
  Matches   19;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTCCTA 22
              |||| || |||||| |||||||
Db        740 TGAGTGATTTCTGGATGTCCTA 761


RESULT 13
US-10-282-122A-23260
; Sequence 23260, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Malone, Cheryl
;  APPLICANT:  Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari
;  APPLICANT:  Zyskind, Judith
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John
;  APPLICANT:  Carr, Grant
;  APPLICANT:  Yamamoto, Robert
;  APPLICANT:  Forsyth, R.
;  APPLICANT:  Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE:  2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
```

```
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 23260
;   LENGTH: 1185
;   TYPE: DNA
;   ORGANISM: Klebsiella pneumoniae
US-10-282-122A-23260

  Query Match             78.2%;  Score 17.2;  DB 12;  Length 1185;
  Best Local Similarity   86.4%;  Pred. No. 80;
  Matches   19;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 TGAGAGAGTTCTGGGTGTCCTA 22
              |||| || |||||| |||||||
Db        743 TGAGCGATTTCTGGATGTCCTA 764


RESULT 14
US-10-104-047-1043/c
; Sequence 1043, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
;  APPLICANT: HELIX RESEARCH INSTITUTE
;  TITLE OF INVENTION: No. US20030236392A1el full length cDNA
;  FILE REFERENCE: H1-A0105
;  CURRENT APPLICATION NUMBER: US/10/104,047
;  CURRENT FILING DATE:  2002-03-25
;  PRIOR APPLICATION NUMBER:
;  PRIOR FILING DATE:
;  NUMBER OF SEQ ID NOS: 4096
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1043
;   LENGTH: 2394
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-104-047-1043

  Query Match             76.4%;  Score 16.8;  DB 15;  Length 2394;
  Best Local Similarity   90.0%;  Pred. No. 1.3e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GAGAGAGTTCTGGGTGTCCT 21
              |||||||||||| ||||| |
Db       1489 GAGAGAGTTCTGAGTGTCAT 1470


RESULT 15
US-10-428-487-11/c
; Sequence 11, Application US/10428487
; Publication No. US20040006780A1
; GENERAL INFORMATION:
;  APPLICANT: RASTELLI, LUCA K.
;  APPLICANT:  GERBER, HANS-PETER
;  TITLE OF INVENTION: VEGF-MODULATED GENES AND METHODS EMPLOYING THEM
;  FILE REFERENCE: 09800080-0103
;  CURRENT APPLICATION NUMBER: US/10/428,487
;  CURRENT FILING DATE:  2003-05-02
;  PRIOR APPLICATION NUMBER: 09/815,153
;  PRIOR FILING DATE: 2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,201
;  PRIOR FILING DATE: 2000-03-22
;  NUMBER OF SEQ ID NOS: 84
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
;   LENGTH: 301692
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-428-487-11

  Query Match             76.4%;  Score 16.8;  DB 15;  Length 301692;
  Best Local Similarity   90.0%;  Pred. No. 1.3e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 AGAGAGTTCTGGGTGTCCTA 22
              ||||| | ||||||||||||
Db      69276 AGAGATTGCTGGGTGTCCTA 69257


Search completed: February 20, 2004, 05:59:27
Job time : 59.8809 secs
```